The Semantic Grid:

will Semantic Web and Grid go hand in hand?

Marije Geldof

European Commission
DG Information Society
Unit “Grid technologies”

June 2004

Contents

1 Introduction

While a decade ago the World Wide Web was still in its infancy, nowadays millions of people around the world could not imagine their lives without it. The WWW has quite rapidly evolved into a vast information, communication and transaction space. Unfortunately this growth goes hand in hand with a decrease in transparency for the user. It can sometimes be difficult to find, access, present and maintain the information required by a wide variety of users. One of the main obstacles is that most information on the Web is made for human interpretation and is not evident for agents browsing the Web. The Semantic Web is an effort to improve the current Web by making Web resources "machine-understandable", because the current Web resources do not respect machine-understandable semantics.

Another promising technology that is still in its infancy is the Grid. With its origins in metacomputing projects to build virtual supercomputers using networked computers in the early 1990s, the vision of the Grid is an infrastructure which delivers computing and data resources seamlessly, transparently and dynamically as and when needed.

Some analogies between the Web and the Grid can be observed; such as they are both frameworks providing the user with something that can be information, communication, computation etc. The Grid is presently mainly studied within scientific communities and the question is how its development will evolve. If the Grid were to go through a rapid evolvement like the Web would this result in a decrease in transparency for the user?

Some people have started to identify the future need for semantics within the Grid and - in order to minimise potential for duplicated effort - they are proposing to bring Grid Computing and the Semantic Web together, eventually leading to the so-called "Semantic Grid". The question is if and how this convergence will evolve in the future? For example: What is the participation of industry and how will it evolve? What are the challenges facing - and what are the critical achievements necessary for achieving - the Semantic Grid? Will the different communities be able to collaborate fruitfully?

This report discusses the current status and future vision on the Semantic Grid. Specifically Chapter 2 presents insight in the Semantic Web and the closely related technology of Web services. Chapter 3 will subsequently explore the Grid and compare it to the Semantic Web. Chapter 4 will then deal with different aspects of the "Semantic Grid". Besides applications and industrial participation, it will discuss potential benefits and the current status of the Semantic Grid. Further it will discuss the necessary cooperation between the different fields, critical issues and the future challenges to be overcome.

The preparation of this report has been helped by the contribution of the following people in the field who have answered some questions about the Semantic Grid that were proposed to them: *Carole Goble*: University of Manchester, *David de Roure*: University of Southampton, *Jim Hendler*: Maryland Information and Network Dynamics Lab, University of Maryland, *Yolanda Gil*: University of Southern California, *Hai Zhuge*: Chinese Academy of Sciences. The following people have indirectly contributed to the questions: *Frank van Harmelen*: Free University Amsterdam, *Bertram Ludaescher*: San Diego Supercomputer Center, *Luc Moreau*: University of Southampton. Their ideas are referred to within the report by their last name in brackets.

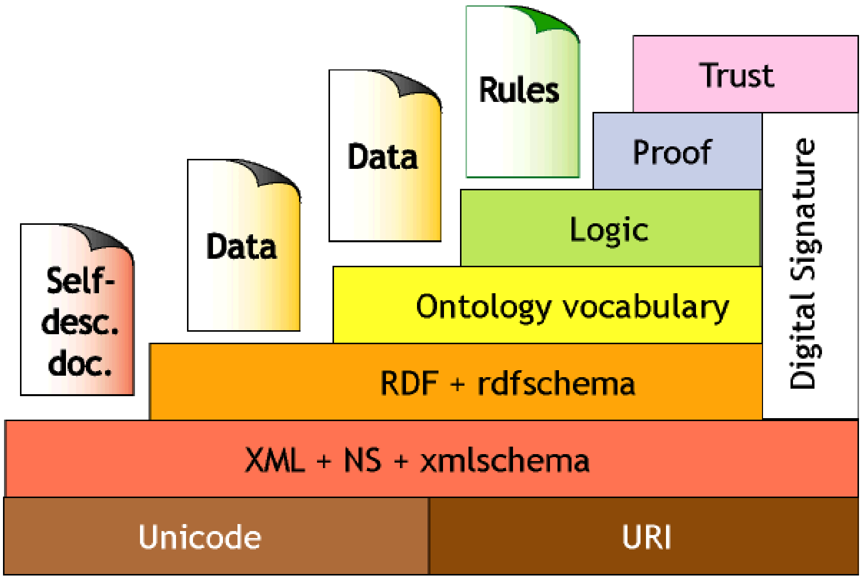
2 The Semantic Web

2.1 What is the Semantic Web?

The Web has brought exciting new possibilities for information access and electronic commerce. It is the Web's simplicity that has fueled its quick uptake and exponential growth, but this same simplicity also seriously hampers its further growth. This is where the Semantic Web comes in:

The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in co-operation. It is the idea of having data on the Web defined and linked in a way that it can be used for more effective discovery, automation, integration and reuse across various applications… data can be shared and processed by automated tools as well as by people [23].

The following graphical representation, originally by Tim Berners-Lee of W3C, shows the proposed layers of the Semantic Web with higher-level languages using the syntax (and semantics) of lower levels.



The Semantic Web promises to make web content machine understandable, allowing agents and applications to access a variety of heterogeneous resources, processing and integrating the content, and producing added value output for users.

2.1.1 Web and Semantic Web – the difference

The Semantic Web is not a separate Web but an extension of the current one. Principally, there are two conceptual differences between the Semantic Web and the Web:

- The Semantic Web is an information space in which the information is expressed in a special machine-targeted language, whereas the Web is an information space that contains information targeted at human consumption expressed in a wide range of natural languages.
- The Semantic Web is a web of formally and semantically interlinked data, whereas the Web is a set of informally interlinked information.

2.1.2 Ontologies

Both Semantic Web and Semantic Grid initiatives build heavily on the utilization of ontologies [12].

An ontology is a formal, explicit specification of a shared conceptualization [10]. It is a system of concepts and their interrelations presented in a machine-understandable format [12]. Using them enables enriching content in a semantically consistent way.

The concept of an ontology is necessary to capture the expressive power that is needed for modeling and reasoning with knowledge. Generally speaking, an ontology determines the extension of terms and the relationship between them. However, in the context of knowledge and web engineering, an ontology is simply a published, more or less agreed, conceptualization of an area of content.

An ontology, apart from the agreed benefits it brings in user navigation, provides common semantics that can be used to improve communication between either humans or computers. Ontologies may be grouped according to their role into the following three areas: to assist in communication between people, to achieve interoperability among computer systems or to improve the process and/or quality of engineering software systems.

The real pick-up of Semantic Web technologies will be in the development of shared ontologies for interoperability (of data, middleware) and metadata in a common data model. This means we will see Semantic Science Webs, where it is not possible to tell if it is a Grid or Semantic Web, and it won't matter (Goble).

2.1.3 Web services

Software programs that can be accessed and executed via the web provide "web services". A service can consist in giving plain information, for example a weather forecast, or it may have an effect in the real world, for instance when booking a flight, ordering a book or transferring money [3]. Thus web services turn the 'static web (of displays)' into a 'web of action' and bring the computer back as a device for computation.

Travel agents for example can offer people the possibility to book complete holiday packages: plane/train/bus tickets, hotels, car rental, excursions, etc. Service providers (airlines, hotel chains etc.) are providing Web services to query their offerings and perform reservations. Credit card companies are also providing services to guarantee payments made by customers. Due to the loosely coupled nature of Web services, the travel agent doesn't need to have a priori agreements with service providers of credit card companies. This allows the travel agent to have access to more services, offering more options to its customers; the credit card companies to offer their services broadly and therefore make their customers happy; and the service providers can offer services broadly and easily and therefore generating more business for themselves.

The great potential of web services puts them at the centre of attention of software developers world-wide, and recent standardization efforts such as UDDI, WSDL and SOAP for advertising, describing and invoking them, aim at providing a more stable platform for their deployment and use. By now Web services seem to have reached a more mature state than the Semantic Web. First attempts have been made to apply Semantic Web technology to Web Services and many interesting future developments are expected at the intersection of Semantic Web and Web Service technology [3].

2.2 Current status of Semantic Web

Researchers and developers world-wide have welcomed the Semantic Web ideas with growing enthusiasm, both in academia and in industry. There is a strong commercial interest in applying Semantic Web technologies and in particular ontologies.

However, the technologies at issue are still at a pre-competitive stage. Their large-scale deployment still requires substantial research. And whether or not the "Semantic Web" is going to repeat the Web's success story of its own accord is still an open question: Its underlying concepts are after all not so easy to grasp and their potential benefits are not so easy to sell [3].

Moreover a critical mass problem has to be solved: adding explicit semantics to content, processes and services does not pay off if no tools are available to make good use of it; developing tools on the other hand, does not pay off if there is little semantically-enriched content to work on [20].

Although expectations are high and most players in the field agree on the enormous potential of these technologies, it is not clear whether commercial interest alone will bring about the momentum necessary for them to become a success [3].

This is a situation where public funding can provide the incentives required to advance research and development up to a point where one can "let things take their course".

Within the Semantic Web research community there appears to be a mutual interest from researchers from the EU and US to cooperate, what resulted in some joint workshops. As part of the worldwide effort that is needed to turn the Semantic Web into a viable global infrastructure for accessing and integrating content and services this cooperation might be of help.

The European Commission has identified the importance of Semantics and it is consequently funding pre-completive research in this area under the 6th framework programme (FP6). Specifically the EU has identified the following research objective for research it will fund:

- *To develop semantic-based and context-aware systems to acquire, organize, process, share and use the knowledge embedded in multimedia content. Research will aim to maximize automation of the complete knowledge lifecycle and achieve semantic interoperability between Web resources and services.*

This has resulted in different Semantic Web related projects. Also within the next work programme there will probably be a new objective for semantics research.

Although the European Commission is funding research programmes to encourage co-operation at European level, still a lot of funding is still managed by national authorities. For many, issues related to the Semantic Web are on their agendas, but mostly implicitly stated. For example Ireland, Germany, Austria and the United Kingdom have national programmes funding Semantic Web related research.

Also in the US several agencies have launched initiatives to develop Semantic Web technologies, sometimes as in Europe, without explicitly referring to the concept itself. The most prominent one is DAML, the DARPA Agent Markup Language programme, designed and brought on its way by James Hendler, an early pioneer of the Semantic Web.

Below you will find a flavor of some of the work going on the Semantic Web area:

- The World Wide Web Consortium (W3C), headed by Tim Berners Lee, is a major international driver behind the Semantic Web.
- In the US further Maryland Information and Network Dynamics Lab has a large Semantic Web Agents Project and Stanford University is involved in the development of the DARPA Agent Markup Language.
- In Europe the Digitial Enterprise Research Institute (DERI) (Galway, Innsbruck), Free University Amsterdam, Universidad Politecnica de Madrid, University of Manchester, University of Karlsruhe are all heavily involved in European Semantic Web initiatives including projects like Ontoweb, Knowledge Web, WonderWeb and Swap.
- The Institute for Learning and Research Technology in Bristol has various Semantic Web research projects.
- Within industry HP laboratories in Bristol has a separate Semantic Web research group and BTexact Technologies in Ipswich is working extensively on the Semantic Web.
- There are several smaller companies, including start-ups, specializing in the Semantic Web.

2.3 Standardization initiatives for the Semantic Web

"The best thing about standards is that there are so many to choose from"

The process of standardizing technology is messy and time-consuming, but it's vitally important. The main standardization initiative for the Semantic Web comes from the World Wide Web Consortium (W3C).

The W3C Semantic Web Activity has been established to serve a leadership role in both the design of specifications and the open, collaborative development of enabling technology. The W3C Semantic Web Activity is additionally focused on effective standards development through the Semantic Web Coordination Group.

Two important Semantic Web standards (called "Recommendations" by W3C) that have received approval of W3C in February 2004 are the Resource Description Framework (RDF) and the Web Ontology Language (OWL). RDF and OWL are standards that provide a framework for asset management, enterprise integration and the sharing and reuse of data on the Web. These standard formats for data sharing span application, enterprise, and community boundaries - all of these different types of 'user' can share the same information, even if they don't share the same software.

On European level the OntoWeb project had the purpose to bring together researchers and industrials, promoting interdisciplinary work and strengthening the European influence on Semantic Web standardisation efforts such as those based on RDF and XML.

3 The Grid

3.1 What is Grid computing?

Usually people don't think about where their electricity is coming from when they plug any electrical equipment in the socket. Grid computing can be seen as analogous to the power grid - a user can have access to computing power and resources on demand and doesn't have to know where the sources are coming from. The concept emerged in the early 1990s as a distributed infrastructure for advanced science and engineering. Meanwhile the focus has shifted further to the interaction of any kind of resource including documents, databases, instruments, archives and people. A more recent definition of the Grid would therefore be: flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources. This is what is often referred to as a "virtual organization": a dynamic collection of individuals, institutions and resources bundled together in order to share resources as they tackle common goals [4]. One of the main purposes of this resource sharing is to support problem solving on a global scale for data-intensive and compute-intensive applications.

Three generations of the Grid can be identified [17]:

- First generation systems involved proprietary solutions for sharing high performance computing resources.
- Second generation systems introduced middleware to cope with scale and heterogeneity, with a focus on large-scale computational power and large volumes of data.
- Third generation systems are adopting a service-oriented approach, are metadata-enabled and may exhibit autonomic features.

It should also be noted there is currently no such single thing as "the Grid", but there are a number of interoperating Grids, a significant proportion of which exist in academic contexts and are maintained by large teams of highly specialized personnel. When talking of "Grids" it is important to make a distinction between Grid middleware and Grid applications. The middleware is used to hide the heterogeneous nature and provide users and applications with a homogeneous and seamless environment by providing a set of standardized interfaces to a variety of services [17]. Grid applications on the other hand represent knowledge and operational know-how of the application domain.

There are two aspects that make the Grid differ from its roots in metacomputing and distributed computing, namely the considerable size of the data handled and the use of sources that are not controlled by the user or its organization. Achieving these aspects requires two things:

- A dynamic way to define and use a computer or data service. This is the goal of the Grid Services.
- A way to describe data and resources in a way that is understandable and usable by the community that is target user. This is the goal of ontologies, part of the Semantic Web.

3.1.1 Benefits of the Grid

There is no potential for a new technology if it doesn't at least offer certain benefits for its users. In short there are two major benefits of grid computing: savings and speed. The benefits can be divided into business benefits and technology benefits. Organizations can therefore experience benefits by implementing grid to achieve results like [12]:

Business benefits

- Accelerate time to results
- Enable collaboration and promote operational flexibility
- Efficiently scale to meet variable business demands
- Increase productivity
- Leverage existing capital investments

Technology benefits

- Infrastructure optimization
- Increased access to data and collaboration
- Resilient, highly available infrastructure

Furthermore the Grid offers some advantages over the Internet [1]:

- There will be a finer granularity of control available, improving security and increasing efficiency.
- There will be a wider variety of resources available over the Grid, for example processing power.
- It will offer a distributed rather than centralized architecture. This means for example that if a resource fails on the Grid it can be replaced by another resource on the Grid.

3.1.2 Grid services

Grid middleware should enable new capabilities to be constructed dynamically and transparently from distributed services. In order to engineer new Grid applications it is desirable to be able to reuse existing software components and information resources and to assemble and coordinate these components in a flexible manner. Partly for this reason the Grid has moved away from a collection of protocols to a service-oriented approach: the Open Grid Services Architecture (OGSA). This unifies Web Services with Grid requirements and techniques.

The Open Grid Services Architecture (OGSA) integrates key Grid technologies with Web services mechanisms to create a distributed system framework based around the Open Grid Services Infrastructure (OGSI). Building on both Grid and Web services technologies; the Open Grid Services Infrastructure (OGSI) defines mechanisms for creating, managing, and exchanging information among entities called Grid services. OGSI also introduces standard factory and registration interfaces for creating and discovering Grid services.

In short, a Grid service is a Web service that conforms to a set of conventions (interfaces and behaviours) that define how a client interacts with a Grid service. These conventions, and other OGSA mechanisms associated with Grid service creation and discovery, provide for the controlled, fault resilient, and secure management of the distributed and often long-lived state that is commonly required in advanced distributed applications.

However web services vendors indicated that while they recognized the importance of OGSI concepts, they would not adopt OGSI as it was then defined. This is why the WS-resource framework was proposed as a refactoring and evolution of OGSI that delivers essentially the same capabilities in a manner that is more in alignment with the Web services community. WSRF retains essentially all of OGSI concepts and introduces only modest changes to OGSI messages and their associated semantics, the changes are primarily syntactic. While there will be certainly some work involved in converting from OGSI to WSRF, it doesn't appear to be substantial.

3.2 Current status of Grid computing

Although Grid computing is gaining a lot of attention within IT industry, Grid computing is not yet a standard product on the ICT market. Some major IT companies and service providers have recently developed and launched grid software products and services. Different Grid research projects have emerged in most of the European countries and also at European level, very often in the frame of the European framework programme for research. Nonetheless a lot of research and development needs to be done to fully realize the concept of Grid computing. The fact that besides the global grid forum there are various national Grid initiatives across the world – from South Korea to Estonia, from Austrialia to Germany from China to the Netherlands – demonstrates that Grid computing has been identified as a serious technology that merits attention on a large scale. Just as an example there are national Grid initiatives in Korea, the Netherlands, Estonia, Australia, Japan, Germany, China etc.

Presently there are only a few real production Grids and not many demonstrations of problems that you could not have been addressed before without Grids. Furthermore current Grid middleware is extremely hard to use for non-specialists users and incomplete. It has provided computational interoperability, but semantic interoperability is now required.
Except for technical constraints non-technical issues have been identified as barriers for industrial uptake of Grid computing, for example unwillingness of people to share computing resources they control with others, even with other colleagues in the same organization [15]. The interest of industry at the moment also mainly seems to stay within the boundaries of the organization. While in academic environments people start to be more used to sharing of resources, industry is a much more competitive environment and still seems to have a long way to go before there will be sharing cross organizational.

The European Commission has identified the importance of Grid research and some of the objectives in the 2003-2004 work programme of the 6th framework programme are as follows

- To expand the potential of the Grid and peer-to-peer approaches to solving complex problems which can not be solved with current technologies in application fields such as, but not limited to, industrial design, engineering and manufacturing, health, genomics and drug design, environment, critical infrastructures, energy, business and finance and new media.
- To overcome present architectural and design limitations hampering the use and wider deployment of computing and knowledge Grids and to enrich its capabilities by including new functionalities required for complex problem solving. This should help the larger uptake of Grid type architectures and extend the concept from computation Grids to knowledge Grids, eventually leading to a "semantic grid".

This has resulted in several Grid projects funded by the European Commission. Also within the next work programme there will probably be a new objective for grid research.

3.3 Standardization initiatives for the Grid

There must be standards for Grid computing that will allow a secure and robust infrastructure to be built. Wider use of Grid computing will depend upon standardization of Grid services. The main standardization initiatives for the Grid come from the Global Grid Forum (GGF) and Organization for the Advancement of Structured Information Standards (OASIS).

GGF is a forum initiated by a community of individuals from industry and research leading the global standardization effort for the Grid. GGF's primary objectives are to promote and support the development, deployment, and implementation of Grid technologies and applications via the creation and documentation of "best practices" - technical specifications, user experiences, and implementation guidelines.

OASIS is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. The consortium produces more Web services standards than any other organization and is responsible for the development of the Web Services Resource Framework (WSRF).

Specifications such as the Open Grid Services Architecture (OGSA) and tools such as those provided by the Globus Toolkit provide the necessary framework. The Globus Project is a multi-institutional research and development effort creating fundamental technologies needed to build grids. And OGSA is a proposed architectural framework for grid computing — a standard that works towards defining clear programming interfaces, management interfaces, naming conventions, directories, and more for the convergence of grid computing and Web services. This architectural specification represents a proposed way to enable grid computing and Web services to merge into one cohesive infrastructure through applications architecture. Although OGSA can still be considered as a developing architecture, recently at GGF11 a step forward was made by the presentation of the first informational version of the specifications for OGSA.

As described in paragraph 3.1.2. first the Open Grid Services Infrastructure (OGSI), followed recently by the WS-resource framework, have been released as standards that allow for the implementation of OGSA services. WSRF was proposed as a refactoring of OGSI concepts to align better with Web services and the current technical ground swell appears to be that WSRF is a better starting point with respect to obtaining wider community acceptance as well as rapid developments of usable tools and Grid applications.

3.4 The Grid and the Semantic Web – similarities and differences

Before we have a look at the convergence of the Semantic Web and the Grid in more detail, let’s have a look at how there two are related. In what way are they similar and how do they differ?

The Grid and Semantic Web both operate in a global distributed and changeable environment [7]. The kind of global and distributed infrastructure needed for both asks for database and information system technologies.

A Grid application might involve large numbers of processes interacting in a coordinated fashion, while typical Web transaction today still only involves a small number of hosts [8]. The Web mainly enables communication, while Grid computing enables full collaboration towards common business goals.

Both need computationally accessible and sharable metadata to support automated information discovery, integration and aggregation. The Semantic Web is a static provider of metadata, while the Grid is a dynamic user of metadata. Grid processes continually appear and disappear and therefore are transient and stateful, while web services persist and are available and stateless [7].

Finally both the Semantic Web and the Grid keep complexity hidden, so multiple users all go through the same experience.

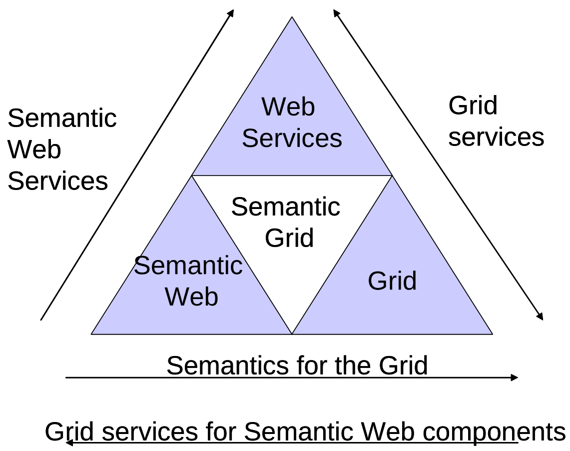
<table>
<tr><th>Semantic Web</th><th>Grid</th></tr>
<tr><td colspan="2" align="center">Operating in global distributed and changeable environment
Computationally accessible and sharable metadata needed
Complexity hidden</td></tr>
<tr><td>Small number of hosts</td><td>Large number of interacting processes</td></tr>
<tr><td>Enables communication</td><td>Enables full collaboration</td></tr>
<tr><td>Static provider of metadata</td><td>Dynamic user of metadata</td></tr>
<tr><td>Persisting web services</td><td>Appearing and disappearing grid services</td></tr>
<tr><td>Stateless web services</td><td>Transient and stateful grid services</td></tr>
</table>

4 Semantic Grid

4.1 What is the Semantic Grid?

Although there is a wide spread feeling that all the fields are interrelated and could benefit from each other, so far there has been little systematic effort in examining at the intersection of P2P, Grid, Semantic Web and Web Services.

One initiative to look at this issue is the so-called 'Semantic Grid', which aims to incorporate the advantages of the Grid, Semantic Web and Web Services. Schematically this can be seen as follows (diagram courtesy of C. Goble):



Nonetheless there doesn't seem to be complete consensus on what the term 'Semantic Grid' exactly means. Some people mainly look at the convergence of the Semantic Web and Grid technologies; others consider it much broader also including techniques like Web Services. There are some who criticise the term 'Semantic Grid' and prefer to talk about a 'Grid with semantics (Kesselman). However despite these different points of views in the end they all aim at adding meaning to the Grid. Therefore it should be noted that in this sense the title of this report can be considered misleading., The report's title was nonetheless chosen as the report aims to address two established communities playing a key role in the Semantic Grid.

Simply speaking the Semantic Grid can be described as an extension of the current Grid in which information and services are given well defined meaning, better enabling computers and people to work in cooperation. It is characterized as an open system in which users, software components and computational resources (all owned by different stakeholders) come and go on a continual basis. It can be seen as a set of services that are offered by entities (which may be software agents). The services are offered under contract, and can be accepted by any of a number of consumers in a marketplace [1].

Currently there is a gap between grid computing endeavors and the vision of Grid computing. This vision sees a Grid in which there is a high degree of easy-to-use and seamless automation and in which there are flexible collaboration and computation on a global scale. To support the full richness of the grid computing vision requires Semantic Web technologies for Grid middleware and applications, i.e. the Semantic Grid. Semantics (knowledge) should be explicitly asserted and used within the Grid software. Aspects of the Semantic Web could be applications of Grid computing, for example in search, datamining, translation and multimedia information retrieval [8]. At the moment Grids are already driven by metadata with the semantics mainly 'buried', although they are there nonetheless.

A distinction can be made between Grid using semantics in order to manage and execute its architectural components (a semantic Grid perspective) and a Grid of Semantics based on knowledge generated by using the Grid; semantics as a means to an end and also as an end itself.

The Semantic Grid concept originally set out to promote the concept of eScience. However the application of the Semantic Grid extends over the boundaries of the scientific community into business, government and industry.

Meanwhile a specific research group has been established within GGF: the GGF Semantic Grid Research Group (SEM-GRD). The purpose of this research group is to help Grid users and developers realize the added value from Semantic Web technologies. In other words bringing Semantic Web technologies and techniques to the Grid.

4.1.1.1 Cognitive/Knowledge/Semantic Grid

Within the field under discussion different terms that seem to be related are used. What is the difference between Semantic Grid, Knowledge Grid and Cognitive Grid? Instead of disuccsiin in detail the difference it might be more useful to talk about their similarities. In the end all these terms deal with adding intelligence to the Grid. The Knowledge Grid is only dealing with sharing knowledge, while the Semantic Grid is dealing with sharing resources in general, not only knowledge, where well defined meaning is added to the resources. The Cognitive Grid is sometimes used interchangeably with the term Semantic Grid and is understood as intelligent management of grid resources. It can be defined as a Grid that incorporates grid services, ontologies and knowledge driven services (Kesselman). The basis for understanding all three terms is the Semantic Web.

4.1.1.2 E-science

The term e-science keeps popping up when searching for information on the Semantic Grid. This term was once introduced by John Taylor, Director General of Research Councils in the UK Office of Science and Technology (OST). "e-Science is about global collaboration in key areas of science and the next generation of infrastructure that will enable it" [11] .

The Grid is a promising technology in enabling this e-science vision, but in reality e-science is broader than the Grid only. The fact that there is a special national UK e-science program funding research in this field where some of the UK initiators of the Semantic Grid have secured funding, might explain why the terms seem to be interlinked.

In the future e-Science will refer to the large-scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large scale computing resources and high performance visualization back to the individual scientists.

4.2 Potential benefits of the Semantic Grid

While the concept of a Semantic Grid can be appealing it is worth exploring the question of why the Grid would need any semantics. First of all if the Grid develops further and take up of Grid technologies accelerates there may be the consequence of reduced transparency for the user as huge amount of resources are not easily interpreted by the human user. By having the interpretation of the resources done by the Grid itself through the use of semantic web technologies, human efforts will be greatly reduced allowing resources to be used more efficiently.

As described by Carole Goble the major benefits that could be delivered are related to heterogeneity and reuse. The real benefits will come from the explicit representation of metadata that can be obtained when there is heterogeneity in data, resources, cross-grid and cross-organisation and when resources need to be reused in different ways and the run of the sessions needs to be modified dynamically.

The reward that will motivate the grid is the prospect of accelerating scientific process and not just scientific computation.

4.3 Applications of the Semantic Grid

For any new technology, of course the best way to make it become a reality and to encourage widespread uptake is to have a revolutionary application that is welcomed by everyone. Creating more interest in the Semantic Grid therefore requires a convincing demonstrator or "killer application" that shows people and industry that it is a promising technology that will deliver a return on investment, be it financial or personal. But what kind of applications could a Semantic Grid provide that would convince for example people in industry that they really need to get involved?

Jim Hendler expects in the short term the first applications to be developed within the biological and life science fields as these communities seem to be the first to realise they will need the Semantic Grid. On the other hand Carole Goble believes that key applications will be in areas where resources have to be brought together, work together and then disband, for example cross organisational insurance settlement. Or applications where the creation of a virtual database or computer or sensor is required that needs dynamic provisions, for example you could then use planning techniques. Alternatively Yolanda Gil proposes composing workflows of services (semi-) automatically by reasoning about the semantics included in service descriptions, the semantics of the message content exchange.

A concrete example could for example be workflows involving multiple data sets and analytical steps, requiring complex semantics for linking things up (Ludaescher).

4.4 Current status of the Semantic Grid

Before looking at the current status of the Semantic Grid activity, it should be noted again that the two fields that are supposed to drive this research namely the Semantic Web and the Grid are both still immature fields. The state of play of the Grid today is reminiscent of the Web some years ago: there is limited deployment, largely driven by enthusiasts within the scientific community, with emerging standards and a degree of commercial uptake. The same might also be said of the Semantic Web.

At the moment Semantic Web technologies such as the Resource Framework (RDF) for metadata representation are increasingly being applied to Grid computing infrastructures and applications, facilitating interoperability and reuse of services, data and tools. However it is agreed that the Semantic Grid still has a long way to go to become a reality. Estimates are made of 3-5 years for concepts and associated technologies to mature and 5 years for the final uptake of the technology (van Harmelen, Goble).

According to Jim Hendler the US Grid community is behind its counterpart in Europe with respect to thinking abut the Grid as a service-based entity, rather than a framework of pipes. He believes that a concerted EU/US partnership is needed that would bring more of the European Semantic Grid thinking to the US research community for the eventual success of this endeavour, as neither side can “do it alone”. He states that cooperation at the funding level between EU and US bodies such as the NSF is be critical to the long-term success of such an initiative.
Currently some research projects that are related to the Semantic Grid can be identified. It shows that at the moment the UK is funding most Semantic Grid related projects within their e-Science program, but also by the joint information systems committee (JISC). Further there are some projects funded by the EU and China is working on a Semantic Grid research plan. The US by comparison has at the time of writing one project funded by the NSF.

4.5 Industrial participation in the Semantic Grid

Industry is still unsure of the Semantic Web and also of Grid other than what can be considered as “cluster computing”. Industry is largely agnostic about semantics and because their “hands are full” with plumbing and data management issues, there is no agenda for longer term issues that require semantics (Gil).
There is at the moment little participation by any industrial organisation in the Semantic Grid as such, but there is a lot of interest in both the Semantic Web and the Grid, but it hasn’t gained enough visibility yet. It is an emerging and new area that has not yet proven its worth that will need governmental funding to “cross the chasm” and get larger industrial buy-in (Hendler).

From the perspective of the “Chinese Semantic Grid Plan” the Semantic Grid needs no involvement of industry and mainly requires fundamental research (Zhuge). Carole Goble and David de Roure on the other hand believe there could be greater industrial involvement. They state that without compelling examples and successes most of industry will back off and instead “wait and see”. The exceptions in Gobles opinion are vendors of Grid computing, but even they have their Grid and semantics departments separated in completely different corporate structures and these should be broken down to deliver a real synergies. De Roure believes that the reuse of resources across collaborators and across time might provide potential savings for companies. Whether this is effective and worth the investment will depend on the business model of the company - reuse might not always be valuable.

4.6 Co-operation between different research communities

The Semantic Grid requires cooperation of different communities, like the one working on Grid technologies and that concentrating on Semantic Web, but also others. The two cited communities have quite different backgrounds. The Grid community originates mainly from the distributed computing community, while the Semantic Web community is largely based on people from knowledge representation and artificial intelligence communities. There are not many people covering both communities and the communities often have completely different skills and agendas.

At the moment there are different communities that explore the composition of services; there are activities occurring in the Global Grid Forum, in the Semantic Web Services, the OWL-group etc. OGSA, SWSI, OWL-S, WMSO, WSD, WS choreography and BPEL are, according to Jim Hendler all doing disjointed work without sufficient collaboration.

A lot of money is being invested in the different research areas but nowhere do they get together. A concern is that reinventing expressive representation languages and delivering efficient algorithms is unnecessary and will take decades (Gil). So what is needed for the Semantic Grid is to bring the different communities together in a single, visible place (Hendler, Gil), for example a research forum and to foster the cooperation by people that will bridge the gap between the communities (Goble, de Roure), like the Semantic Web once made Artificial Intelligence and database people talk together. There are already some early pioneers that are trying to bridge the gap like David de Roure, Yolanda Gil, James Hendler and Carole Goble.
According to Carole Goble the European Commission does not yet provide framework for this. The funding for Grid and Semantic Web research is divided over different units without much cooperation among them. She proposes joint programmes across the DG to foster the Semantic Grid for example by providing introductory or training days for the different communities.

Still the question remains why the communities would be interested in cooperating, what is in for them? According to David de Roure the benefits of using semantic web technologies in grid applications are clear, which gives interest and engagement from the grid community. There is however less participation from the semantic web community, partly because this community is still a bit dispersed itself. One part of the community is mainly focused on the ‘pure’ semantic part and is not interested in the Grid, while another part of the community is more pragmatic and sees the role of the Grid.

Besides of course a possible reinvention of certain wheels and the possible needed effort to bring the technologies together in a later stage, when they have evolved in their own directions, Grid computing can benefit from the Semantic Web fabric and services for the management of its semantics. The Semantic Web on the other hand can benefit from the application pull provided by the Grid and the Grid infrastructure itself [7].

It should be remembered at all times that when talking about the convergence of the Semantic Web and the Grid; this should be seen as a two-way traffic. The question is not only what the Semantic Web can do for the Grid, but also what the Grid can do for the Semantic Web (Goble).

4.7 Critical issues facing the Semantic Grid

Although a lot of work has been undertaken, both the Grid and the Semantic Web still are not a reality, they can still be considered to be in an experimental stage, especially Grid middleware that is still immature in many key respects. Therefore talking about “convergence” really means talking about the convergence of two concepts to a new concept. The question is how the evolvement of both fields and the convergence will interact. Will the convergence stimulate or suppress the evolvement of both fields? Will it lead the development of the fields in a direction different to that which would be pursued if they were to develop separately? And what if one of the fields in the end does not succeed and is taken over by a new more promising technique? What will the consequences be for the Semantic Grid?
The answers from some experts in the field make clear that some people still tend to stick to their own point of view and stay on their “side of the bridge”. Therefore one of the first steps

would be to facilitate people of different fields leaving their own focus and to think more broadly than what they themselves are interested in.
For the Grid and also the Semantic Grid really to become a reality people have to start thinking beyond the boundaries of their organisation and have to get used to the idea of sharing with almost everyone. This means that the people developing and investing in the Semantic Grid are probably the first to give a good example. They should consider how to cooperate with anyone and not just how to be the most competitive among others.
Technologies must not be developed for the sake of developing technologies. They should respond to real needs and they will be successful (commercially and otherwise) only if they do so. On the other hand it is worth noting that the current Web technology grew out of a research environment, but also has turned into big business very rapidly.

Some worries related to the Semantic Grid:
- There are too many standards and standard bodies. In some respects this results in defeating the purpose of standardisation.
- Low availability of production examples or “killer apps” (something you could not have done before) available (Goble, de Roure)
- Deployment, research, development, applications, standardization, commercialization are all happening at once.
- Demonstrated scalability capabilities are needed.
- OGSA is still in its infancy. Premature standardization involves the risk of “half-baked” solutions.
- The current middleware is hard to use and incomplete and far from “invisible”)

4.8 Challenges to be overcome

As outlined above there are still a lot of things to happen in the Semantic Grid area, before it can deliver a wider impact. Therefore what are the challenges for this field?
First of all one of the first challenges is to get the people from the different fields first of all talk to each other. Some people state it as having the plumbers talk with the conceptual people (Goble, van Harmelen). Once they are talking to each other the next challenge would be to have them working together and not have them working for their own ends. And the cooperation should have a worldwide focus. This will ask for high entry cost from both sides, as one has to invest a lot of energy in understanding a new area. It is easier for researchers to continue working on the particular areas of research rather than embarking in expanding their horizons to a very different discipline (Gil).
From the semantical point of view there are several technological breakthroughs necessary (van Harmelen):
- Scalability, performance, robustness, reliability
- Privacy, security, access-rights
- Dealing with dynamic information, state, QoS, policies
- Lowering entry point of uptake of semantic technologies
- Some (partial) answer to ontology mapping/integration

From the grid point of view the following change is necessary (van Harmelen):
- Move from fixed-pipeline processes to dynamic compositions

Furthermore the Grid is a changing technology, which Semantic Web technologies would have to reflect. Semantic Web services should be brought into Grid environments, making them Grid services, just like other Grid services (Goble).
Of course the Semantic Web and Grid communities are the most obvious ones for a contribution to the Semantic Grid. But there are also other communities that can make a

contribution, so besides the Semantic Web and Grid communities other relevant communities should be involved.
In practical terms the development of a Semantic Grid first of all asks for stability in both the Grid and the Semantic Web to ensure environments are robust and readily usable (Goble). The Semantic Grid can only become a reality if high quality of service is offered to users and it facilitates delivery of applications at all levels of the Grid fabric. This can be fostered by community-based standards, for example the GGF can play an important role there.

Another important challenge, maybe even the most crucial one in the end, relates to the uptake of the Semantic Grid. Benefits of the Semantic Grid will have to be shown to a larger public in order to secure widespread acceptance of this new technology, on top of identifying and addressing industrial need for the Semantic Grid.
Since it is too early to expect the corporate world to accept the current solutions, a critical mass needs to be generated. A first step in addressing this challenge would therefore be to bridge the gap from "research" to practice with the support of governmental funding (Hendler). Furthermore, the research agenda for longer term issues should be determined and David de Roure pleas for bigger projects and overarching activities, that can make the Semantic Grid more visible than by small projects.

A good demonstrator should be developed that can clearly show the return on investment (de Roure). Furthermore some "low hanging fruit" should be provided to make the Semantic Grid more visible, for example manual/semi-automatic configuration of Grid services currently requires very significant amounts of specialized knowledge (van Harmelen).

The immediate adoption of the technologies that are available now should be encouraged. Carole Goble expects there will be real demonstrators within 2 years in a closed scientific community, most likely Life Sciences.
The ideal scenario would be if the Semantic Grid were to allow something to be preformed that wouldn't have happened or have been possible without the Semantic Grid, such as facilitating an international research project which bridges disciplines may create new disciplines. It is, for example, the expectation of David de Roure expects this could be possible two years from now.

More listings of research challenges of the Semantic Grid can be found in [18] and [15].

5 Conclusions

Although there are promising perspectives we are still far away from the Semantic Grid becoming a reality; the future will reveal much about how the Grid and Semantic Web will finally converge.

The biggest challenge to be overcome appears to be of a non-technical nature: having people from different fields talk, work and evolve together. Another challenge is that overcoming reluctance of industry and business to fully engage Semantic Grid technologies given many in industry and business remain somewhat unsure of the Semantic Web and the benefits of the Grid (other than 'cluster computing'). Nonetheless commercial interest alone will probably not drive the Semantic Grid to become a success.

This is a situation where public funding can play a role in fostering the research up to a point where it is taken over by users outside of the research community, be it in business or the general public. At the moment most of the Semantic Grid related funding seems to come from the UK and to a less extent from the EU via research programmes overseen by the European Commission. Public institutions, like national or international programmes, should decide to take a fostering role in the future of Semantic Grid development. For example by initiatives that introduce the different people to each other and by encouraging and funding research projects that bridge the various fields. However the separation of the different fields is often also reflected in the public institutions themselves, for example by separated units in international institutions covering the Semantic Web and Grid fields. Consequently in order to successfully stimulate the development of the Semantic Grid, first of all a better cooperation within institutions themselves should be considered.

The question that remains unanswered is how the two fields will go together. Will there be the need for a separate Semantic Grid community or will good collaboration between different communities suffice? Activities over the coming years will reveal the answer to this fundamental question.

6 Acknowledgements

The author wishes to thank everyone that provided input to this report. Special thanks to Carole Goble, Jim Hendler, Yolanda Gil, David de Roure and Hai Zhuge for their contribution by providing answers to some questions about the Semantic Grid. Furthermore to Annalisa Bogliolo and Eoghan O'Neill for their assistance.

7 Bibliography

1. Adcock, S., How does the grid extend the internet, and what is the future vision for this development?, 2000 http://mms.ecs.soton.ac.uk/papers/27.pdf
2. Chen L., Shadbolt N.R., Tao F., Puleston C., Goble C., Cox S.J. (2003) Exploiting Semantics for e-Science on the Semantic Grid. *Web Intelligence (WI2003) workshop on Knowledge Grid and Grid Intelligence*, 122-132.
3. Ding, Y., Fensel, D., and Stork, H. (2003). The Semantic Web: from Concept to Percept. *Austrian Artificial Intelligence Journal*, 21 (4).
4. Foster, I., Kesselman, C. Tuecke, S. (2001), The anatomy of the Grid: Enabling scalable virtual organizations, *International Journal of Supercomputer Applications*, 15(3).

5. Foster, I., Kesselman, C., (1998) *The Grid: blueprint for a new computing infrastructure*, Morgan Kaufmann, July 1998. J. UCS 8(9): 848-.
6. Fox, G., (2003). Data and metadata on the semantic Grid. *Computing in Science & Engineering*, 5(5), 76-78
7. Goble, C.A. and De Roure, D. (2002) The Grid: an application of the semantic web. *ACM SIGMOD Record* 31(4): 65-70.
8. Goble, C. and De Roure, D. (2002). The Semantic Web and Grid Computing, in Kashyap, V. Ed. *Real World Semantic Web Applications*, IOS Press.
9. Goble, C.A., De Roure, D., Shadbolt, and Fernandes, A.A.A. (2004) Enhancing Services and Applications with Knowledge and Semantics in Foster, I. and Kesselman, C., Eds. *The Grid 2: Blueprint for a New Computing Infrastructure (2nd edition)*, Morgan-Kaufmann.
10. Gruber, T.R. (1993), A translation approach to portable ontology specifications, *Knowledge Acquisition*, 5, 199-220.
11. Hey, T. and Trefethen, A. (2003) e-Science and its implications. *Philosophical Transactions of the Royal Society*, 361(1809), 1809-1825.
12. IBM, http://www-1.ibm.com/partnerworld/pwhome.nsf/weblook/pat_sol_grid_benefits.html
13. Korda, N,. Semantic Grid: Needs, Ideas, Prospects, www.ejewish.info/reka/seminar_ppt/article_korda.doc
14. Newhouse, S., Mayer, S., Furmento, S., McGough, S., Stanton, J., and Darlington, J. (2002) Laying the Foundations for the Semantic Grid. *Proceedings of AISB Workshop on AI and Grid Computing*.
15. Platform Computing Inc. The Non-Technical Barriers to Implementing Shared Computing in a Commercial Environment, Market Study Report, March 2003. www.platform.com/barriers
16. De Roure, D., Jennings, N.R. and Shadbolt, N. R. (2003) The Semantic Grid: A future e-Science infrastructure, in Berman, F., Fox, G. and Hey, A. J. G., Eds. *Grid Computing - Making the Global Infrastructure a Reality*, 437-470. John Wiley and Sons Ltd.
17. De Roure, D., Baker, M.A., Jennings, N.R. and Shadbolt, N.R. (2003). The evolution of the Grid, in Berman, F., Fox, G. and Hey, A.J.G., Eds. *Grid computing – making the global infrastructure a reality*, 65-100, John Wiley and Sons Ltd.
18. De Roure, D. and Hendler, J.A. (2004). E-Science: the Grid and the Semantic Web. *IEEE Intelligent Systems*, 19(1), 65-71.
19. Siddiqi, J., Akhgar, B., and Naderi, M. (2003). Towards a Grid Enabled Knowledge Management Services, http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/073.pdf
20. Stork, H. (2002). Webs, Grids and Knowledge Spaces: Programmes, Projects and Prospects.
21. Tao, F, Cox, S.J, Chen, L, Shadbolt, N.R, Xu, F, Puleston, C, Goble, C, and Song,W.(2003). Towards the Semantic Grid: Enriching Content for Management and Reuse. *Proceedings of UK e-Science All Hands Meeting*, 695-702.
22. Tao, F., Chen, L., Shadbolt, N. R., Pound, G. and Cox, S. J. (2003) Towards the Semantic Grid: Putting Knowledge to Work in Design Optimisation. In *Proceedings of The 3rd International Conference on Knowledge Management (I-KNOW '03)*, 555-566.
23. W3C Semantic Web Activity Statement, http://www.w3.org/2001/sw/Activity/
24. Zhuge, H., (2004). Semantics, Resource and Grid. *Future Generation Computer Systems* 20(1), 1-5.