TextGrid - a modular platform for collaborative textual editing

Andreas Aschenbrenner[1], Peter Gietz[2], Marc Wilhelm Küster[3], Christoph Ludwig[3] and Heike Neuroth[1]

[1] Goettingen State and University Library, Germany
{aschenbrenner, neuroth} at sub.uni-goettingen.de
[2] DAASI International Ltd., Tübingen, Germany
{peter.gietz} at daasi.de
[3] Fachhochschule Worms, Germany
{kuester, ludwig} at fh-worms.de

Abstract. TextGrid equips textual criticism with a fundamental infrastructure and a powerful set of software tools based on the evaluation of existing solutions and embracing the grid paradigm. Its technology and services effectively support text scholars in their daily working processes. With its distinctive features and being a native grid project for the humanities, TextGrid roots in domain sciences including textual scholarship, the Digital Library as well as the e-Science communities, and it aims to span the continuum between them. After its launch in early 2006, TextGrid has three years to achieve its goals and make e-Humanities a reality. During all this time the project will be open and invite participation, to attain a generic software platform that can be used and re-used in multiple contexts.

Keywords: e-Science, e-Humanities, digital libraries, grid, arts, humanities, cultural heritage

1 Introduction

Textual criticism researches the complex processes in the genesis of texts and the contexts of literature. Despite modern information technology and a clear thrust towards collaboration, text scientists still mostly work in local systems and project-oriented applications. Current initiatives lack integration with existing text corpora and remain unconnected to online resources such as dictionaries, lexica, secondary literature and tools. However, this integration and interconnection bears a wealth of opportunities. Integrated tools that satisfy the specific requirements of textual scholarship could transform all relevant processes - including annotation, edition, text processing, and publication. Working towards this vision, TextGrid aims at building a virtual workbench based on grid technologies and e-Science methodology.

A grid-enabled architecture for TextGrid promises to achieve this integration in a dynamic environment on various levels. Past and current initiatives for digitising and ingesting texts already accrued considerable data volumes exceeding multiple terabytes. Storage grid technologies are capable of handling these data volumes effectively. Also the geographical scattering of the community and the scattering of resources and tools suggest that textual scholarship is ripe for grid technologies.

TextGrid[1] is a three years effort, and with its launch in February 2006 it has only just started. The project is publicly funded by the German Federal Ministry of Education and Research (BMBF) as part of the ongoing global movement towards grid infrastructures and the convergence of e-Science technologies in general. Project partners[2] include both, technologists and domain experts from six German higher education institutions and two commercial partners.

The upcoming subsection briefly presents D-Grid, the German grid initiative and the umbrella organisation of TextGrid. The TextGrid vision - inter-weaving existing repositories and tools, as well as providing a platform for connecting experts and enabling related initiatives - is further described in the next sub-chapter. The TextGrid architecture and other technical issues are outlined in section 2.

1.1 D-Grid

Numerous efforts to establish large-scale e-Science infrastructures have been launched on a national or international scale, including the UK e-Science Programme[3] and the US Cyberinfrastructure program[4]. In 2004, the German Federal Ministry for Research and Education (BMBF) triggered the German e-Science initiative[5] as a long-term strategic grid research and development initiative. The D-Grid[6] programme is part of the German e-Science initiative, and aims to establish a generic infrastructure based on grid technologies for other e-Science projects to plug into. D-Grid combines over 100 scientific institutions with public funding of 100 million Euros over a period of 5 years.

The first D-Grid call[7] inaugurated a grid infrastructure project and six community projects[8] for a three-year phase, which started late 2005. [1] The D-Grid Integration Project (DGI)[9] provides basic resources and services for the basic grid infrastructure. Community Grids build upon this fundament and establish domain specific applications in areas including high-energy physics, astrophysics, medicine and life sciences, earth sciences (e.g. climate), and engineering sciences. Textual criticism provides an ideal scenario for transferring e-Science methodologies and grid technologies into a field of the humanities. It is for this reason that TextGrid is funded

1 TextGrid, a community grid for the humanities. www.textgrid.de

2 TextGrid Partners include the State and University Library (SUB) Goettingen, universities in Darmstadt, Trier, Worms and Würzburg, the Institute for the German Language (IDS) Mannheim, as well as the two commercial partners DAASI International and Saphor.

3 UK e-Science Programme, http://www.rcuk.ac.uk/e-Science/

4 US Cyberinfrastructure programme, http://www.nsf.gov/od/oci/reports/toc.jsp

5 German e-Science initiative, http://www.bmbf.de/de/298.php

6 D-Grid, http://www.d-grid.de/

7 The first D-Grid Call, August 2004 (in German) - http://www.pt-it.de/in/e-Science/docs/E-science_Call04_PTIN.pdf

8 Projects in the first D-Grid call include the D-Grid Integration project (DGI, http://dgi.d-grid.de/); AstroGrid-D in astronomy (www.gac-grid.de); C3-Grid for climate research (www.c3grid.de); HEP-Grid for high energy physics (www.hepcg.org); InGrid for engineering research (www.ingrid-info.de); MediGrid for medical research (www.medigrid.de); and TextGrid for humanities (www.textgrid.de).

9 D-Grid Infrastructure Project (DGI), http://dgi.d-grid.de/

as part of the first D-Grid call, being the first project in the German e-Science landscape to bridge the gap between current grid application areas and the humanities.

1.2 Mission

Scholars in text criticism and text editing conduct a myriad of philological, linguistic, and historic research activities on literary and linguistic texts. While the bulk of the work in the analysis, annotation, edition, and publication of textual data requires human input, semi-/automatic tools assisting the activities play a key role for attaining efficiency and quality. Text researchers thus perform different tasks and apply various tools to achieve their goals, using steps of automatic processing intermitted by human review and refinement. Mostly, these one-step tasks such as the collation of texts or lemmatising have several approaches for automatic processing and they need to be fitted to the specific languages and application domains.

Despite their different perspectives, expertises and goals, disciplines in textual scholarship - including philology, linguistics, and lexicography - share some commonalities on a tool-level and can effectively be combined in a single platform. Particularly when focusing on historic texts, digitisation, transcription and encoding of the original manuscripts are common requirements. Once a text is encoded and subsequently augmented with metadata, initiatives annotate the material according to the needs of their research focus. For instance, words may be annotated with regards to their grammatical form (part-of-speech tagging), they may be lemmatised. Countless other forms of annotation are equally conceivable. In philology, the different versions of a text and the contributions of authors/parties over time are carefully noted in order to research the genesis of a text. All these activities involve iterative steps of annotation and analysis. Eventually the annotated textual objects or a whole corpus of texts may be made available for other scientific projects, an analytic publication embedding parts of the annotated text and digitisations may present the derived results, and/or a novel edition of a historic text may be published.

Text researchers created a multiplicity of tools to this end. Indeed, computer assisted text processing has been employed for several decades. TUSTEP[10], for instance, a widely employed application for scholarly processing of textual data, has its origins in the 1960s. Humanities computing [2, 3] and related areas are active fields of research. However, the tools created by relevant initiatives are often local and tailored to specific requirements. They may be outstanding pieces of technology, yet lack the interoperability to be embedded in a chain of tools. Furthermore, they often embody software development paradigms that are no longer popular.

Once research activities span a whole corpus containing numerous texts, weaving the diverse tools together and performing the various tasks in the respective environments becomes an almost intractable feat. One such activity may be identifying and analysing the context of specific phrases within all the works of a particular author, or searching where a particular word appeared first in all available texts of the 17th and 18th century. The integration of the various tasks and tools in a homogeneous framework hence enables the streamlining of existing research processes and may lead to totally new approaches.

10 TUSTEP, http://www.zdv.uni-tuebingen.de/tustep/

This is exactly where TextGrid ties in. TextGrid aims to establish a generic platform with a set of key tools for scholarly text processing, which is extensible for new tools to be integrated and open for reusability. The tools to be included in TextGrid fulfil a range of tasks in processing, analysing, indexing, annotating, editing and publishing textual data for academic research. The project reviews existing tools for adaptation, re-develops them where necessary or useful, and aims to include current research regarding information engineering and semantic technologies. Conceivable tools for the latter include automatic semantic annotation through linkage with dictionaries or ontologies similar to WordNet [4].

Beyond the integration of individual tools, TextGrid offers the possibility to recombine tools and to reuse results from one task in another processing step coordinated by a workflow management system. Network effects turn the TextGrid workbench into more than a mere assembly of individual tools. Interoperability within TextGrid and to external initiatives is achieved through open interfaces and the use of open standards. Standards in TextGrid include the format for literary and linguistic texts by the Text Encoding Initiative (TEI)[11], as well as relevant Web Service and Grid standards[12] on the technical side. [5]

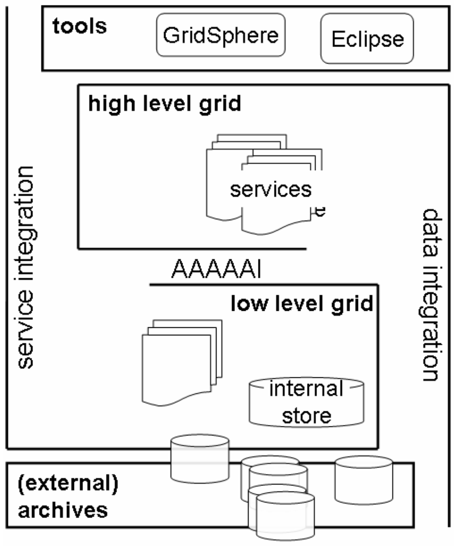
Apart from the basic infrastructure, the resources themselves are of course the most important asset in TextGrid. Initial content is provided by project members, [13] and in the future just any archive with relevant scientific texts can be integrated. Initially, only open content is included into the TextGrid, yet questions pertaining to rights management of licensed material may be reviewed in the course of the project.

2 TextGrid Architecture

With the TextGrid vision clearly laid out, the architecture has to satisfy a number of resulting requirements. The architecture needs to enable tight integration and links on all levels. Text documents are embedded in their context of other documents and linked to dictionaries, annotation services, and other resources. The interconnection of dispersed archives is key in the realisation of the TextGrid vision. Services and tools that can be "re-mixed" dynamically [6] will be integrated in the overall TextGrid workbench successively. Besides inter-connecting data resources and tools, TextGrid aims to link its human users and foster collaboration. Furthermore, TextGrid aims to become the linking factor in a web of initiatives in textual scholarship. Thus, an open, expandable and interoperable architecture is paramount in achieving the TextGrid vision.

These key requirements for achieving the TextGrid vision suggest a loosely-coupled system as basic architectural paradigm. Loosely-coupled systems such as service-oriented architectures, peer-to-peer and grid computing have been suggested as powerful architectures for digital libraries previously. [7] From these architecture types, peer-to-peer systems focus particularly on ad-hoc connections designed to equally integrate every participant in the network. However, they are mostly not only

11 Text Encoding Initiative (TEI), http://www.tei-c.org/
12 TextGrid builds fundamentally on the Web Services Resource Framework (WSRF), which embeds XML, WSDL, SOAP and other relevant W3C Web Service standards.
13 See section 2.1 for exemplary text corpora to be included in the TextGrid.



TextGrid architectural framework

loosely coupled but also loosely integrated, allowing a high level of heterogeneity with only few core operations available for diverse resources. For the TextGrid to emerge as laid out in the vision and the requirements above, a tight integration of the text archives is essential, to enable quality services spanning all resources plugged into the platform. Therefore, another loosely-coupled architecture type with a higher integrative factor is most suitable - grids.

Grids are hardware and software infrastructures, which provide "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations." [8] Grid middleware software including Globus[14], gLite[15], and Unicore[16] provide software toolkits for establishing such an infrastructure. The toolkits include modules for job scheduling, data management and information services, security and rights management, monitoring, and others. This is sometimes called an Infrastructure for Authentication, Authorisation, Accounting, Audit, and Access (here referred to as AAAAAI). [9]

The distinctive features of the TextGrid architecture can be characterised on two different levels: a low level grid infrastructure that provides resource allocation for storage and computing, and a higher level service-oriented grid providing general tools and services for text processing.

As displayed in the TextGrid architectural framework graph, the low level grid interfaces to the archives holding text assets in a distributed and potentially heterogeneous environment. Despite the heterogeneous sources, data integration needs to be achieved for collection-spanning functionality to emerge. The functionalities tailored to the requirements of textual scholarship are embedded in the service layer of the high level grid and made accessible through stand-alone (Eclipse

14 Globus Toolkit, http://www.globus.org/toolkit

15 gLite, Lightweight Middleware for Grid Computing. http://glite.web.cern.ch/glite

16 UNICORE - UNiform Interface to COmputing REsources. http://www.unicore.org/

based) tools, possibly augmented by web-based tools (employing the GridSphere portal framework). The grid middleware components (AAAAAI) are presented in the middle of the graph, and they provide the overarching infrastructure.

One of these components covers the authentication and authorisation infrastructure (AAI). AAI software in existing grid middleware includes the Virtual Organisation Membership Service (VOMS) [10]. In the libraries domain, however, the Internet2 software Shibboleth[17] is fairly wide spread and its general adoption in German higher educational organisations can be foreseen. [11, 12] TextGrid thus needs to interface to Shibboleth in order to address its target community, even though Shibboleth is not native grid software. Others share this requirement, and projects on integrating Shibboleth into grid middleware such as GridShib are ongoing [13]. This issue has also been raised within D-Grid , and TextGrid will cooperate with various partners to achieve this integration.

2.1 Low level grid infrastructure and storage grid

TextGrid's vision is to offer an open platform for services as well as resources. In this vision, the valuable resources created in the course of isolated projects are combined to a single virtual space. A distributed storage grid enables this virtual space, based on Grid standards like WSRF as well as on interfaces and interoperability standards defined in cooperation with collaborating text archives.

TextGrid partners - specifically Würzburg, Trier, and Göttingen - have an excellent record in the digitisation of historic texts and their scholarly edition. Their archives contain valuable textual data in various formats that will be integrated in the TextGrid. As an initial test-bed, two collections from Würzburg are used. The "Dictionary of German Language" by Joachim Heinrich Campe from the early 19th century is a model for a highly structured document. Its digitised images of the 6000-page dictionary come up to a total volume of 600 gigabytes of data. A transcription into XML mark-up is currently being created. The other test-bed collection is the historical-critical edition of works by Jean-Paul[18], which includes 4000 pages of handwritten pages.

Starting from the Würzburg model collections, the TextGrid is designed to be extensible to other TextGrid partners and to just anybody interested. Various questions pertain to the distributed management of the data. To enable the seamless integration of text collections into TextGrid, some data may need to be extracted from the original archives, such as comprehensive index data for enabling the community-specific requirements for complex query and retrieval functionalities.

A variety of grid middleware data management components are available for handling the basic tasks involved, such as integration of heterogeneous storage architectures, data replication, or brokering of virtual files. The Globus middleware [14] offers a toolbox for grid data management[19], the D-Grid Integration Platform supports for example dCache/SRM[20], and there are other relevant packages.

---

17 Shibboleth, http://shibboleth.internet2.edu
18 Arbeitsstelle Jean-Paul-Edition, Würzburg University http://www.uni-wuerzburg.de/germanistik/neu/jp-arbeitsstelle/
19 Globus Components for Grid Data Management, http://www.globus.org/grid_software/data/
20 dCache/SRM, http://www.dcache.org/

However, integration in TextGrid goes beyond the mere connection of possibly heterogeneous archives. A clear convergence in terms of data formats, metadata, and interfaces is needed to allow for the transparent processing of all assets in the TextGrid across repository borders. Where adaptive services cannot do the job of integrating heterogeneous environments, compliance to standard formats may be requested. TEI as a possible standard data format for the editing components in TextGrid has already been mentioned above.

When including a collection into the TextGrid, integration on a semantic level obviously demands a set of operations besides the installation of the storage grid middleware, ranging from data export for TextGrid-wide indexing, to data conversion for standards compliance. The TextGrid community hence needs to account and provide for adequate migration paths and other relevant tools to encourage and support the inclusion of text collections.

One key challenge for establishing TextGrid is the efficient and comprehensive brokerage of objects. Obviously, all objects in the TextGrid should be retrievable, and smoothly embedded in automatic services and user interactions. For example a user may want to conduct a processing step on all works by a specific author within a few clicks. At the same time, there may be a multiplicity of variants of one specific object in the TextGrid. For instance a digitised version of the novel "Hesperus" by Jean Paul may be available in TIFF format, automatically generated and post-processed plain text, versions in TEI mark-up, and various intermediate formats by different editors. So which one to choose for which activity? How can an editor be made aware of already existing artefacts to avoid duplication of work? - This raises a myriad of questions that span various TextGrid components, from low-level forms of representation and metadata to web service definitions and the user interface. There are numerous conceivable solutions to this, some of which demand a higher user effort and others which fall short in comprehensiveness. In the ideal case, a level of machine comprehension could be induced, which conducts some of the reasoning in place of the user. The optimal and most suitable solution remains to be fathomed.

All in all, TextGrid aims for tight integration of text archives on a data as well as a semantic level. This puts several requirements to archives at inclusion time and during ongoing operation, which should be kept as low as possible. In the general case, however, text archives with their painstakingly assorted collections hardly expect to be "hot-plugged" into the TextGrid. Their striving for quality and comprehensiveness concurs nicely with the TextGrid goal of tight integration.

2.2 Higher level service-oriented service grid

Besides the resources, TextGrid's most salient feature is its tools and services. The goal is an open service layer, where any initiative can add their services or re-mix existing services for attaining the specific functionality needed in their research. Thus, the project aims to approach external parties early on, in order to collaboratively define interfaces and interoperability standards.

For several years there has been a clear trend of convergence between grid and web service technologies. [15] The Open Grid Services Architecture (OGSA) [16] has been an important step on this route towards a grid system architecture based on web service concepts and technologies. The trend culminated in the move towards the

Web Services Resource Framework (WSRF)[21] by major grid middleware packages including Globus 4 and gLite.

As mentioned above, the objective is a modular and extensible toolkit of services and tools. Services fulfil functionalities such as word stemming or text comparison. These atomic tasks are coordinated in a workflow management component and integrated into tools tailored to the needs of a specific research initiative. An editor may, for instance, analyse a set of historic texts with dedicated services using dictionaries from that time, and contrast the results in a single view.

Tools come in different flavours. Firstly, web-based tools enable basic tasks such as viewing and querying TextGrid assets. Scholarly editing, however, is an extensive and interactive undertaking with phases that require mainly human creativity. Therefore, the interactive activities in scholarly editing can be done in a locally installed user interface based on Eclipse[22]. Eventually, editors may choose to work in batch mode for processing large text quantities with a predefined workflow. These batch processes are supported via direct access to services through programming interfaces, or by the workflow module and its interactive user interface, the workflow editor.

The workflow management module takes care of assembling and triggering web services in the case of batch jobs or any other more mechanic process that can be executed automatically. Despite their largely automatic nature, these processes can be fairly complex with branches and composite tasks. Moreover, the workflow manager will have to integrate non-automatic tasks that require user interaction.

There are of course many more architectural components and challenges TextGrid faces. However, this brief account of the most pronounced and distinct issues conveys a first characterisation of what TextGrid is about from a technical perspective.

3 Conclusions

TextGrid is a native grid project. Both, the storage grid and the service grid modules will be based on existing grid technology. However, many TextGrid partners are domain experts in textual scholarship and have a digital library tradition. As an initiative that spans both communities, it has lots to learn from and contribute to them. Moreover, TextGrid is one of the early e-Humanities projects in Germany. Initiatives such as the e-Science focus of the Arts and Humanities Data Service (AHDS) in the UK[23]; the University of California Cyberinfrastructure for Humanities, Arts and Social Sciences[24]; or the Humanities and Social Sciences Research Group at the Global Grid Forum[25] indicate the benefits of and trend towards applying grid technologies in the humanities, arts, and social sciences.

However, e-Science goes beyond grid technologies and loosely-coupled architectures. e-Science is also about collaboration, exchange of tools, resources, and knowledge, and about linking people as well as linking resources. The Social

21 Web Services Resource Framework (WSRF), http://www.oasis-open.org/committees/wsrf
22 Eclipse, http://www.eclipse.org/
23 Arts and Humanities Data Service (AHDS), e-Science focus - http://ahds.ac.uk/e-science/
24 University of California Cyberinfrastructure for Humanities, Arts and Social Sciences http://www.uchri.org/main.php?page_id=154
25 Humanities & Social Sciences Research Group, http://forge.gridforum.org/projects/hass-rg

Software movement [17] has first indicated the vast opportunities that remain to be explored in this respect. TextGrid aspires to transfer not only technologies but also e-Science concepts and attitude into textual criticism.

One key research area for the TextGrid project will be the take-up of semantic technologies to further the development of the Semantic Grid. [18, 19] It has a clear target of attaining a tightly integrated platform, which goes beyond the management of distributed resources and rather interweaves all its assets on a semantic level. In addition to this the project research into the provision of semantic services, including the application of statistical methods and ontologies for automatic text analysis.

Apart from technology research, the project aims to become a self-sustainable infrastructure, active beyond the project duration. The involvement of its broad target audience from the very outset is a key part in establishing an active user community.

As the hinge between distributed text archives, TextGrid also has a stake in fostering interoperability between them, and promoting technological and organisational activities for ensuring the long-term preservation of their assets. Long-term preservation challenges are particularly pronounced in a context of historic documents, created by scientific initiatives that may be unaware of preservation practices and held in heterogeneous repositories that may or may not allow preservation functionalities to be retrofitted. Generally speaking, preservation must be an issue for any grid initiative involving valuable digital objects. Bearing all this in mind, TextGrid's stake in digital preservation goes beyond one of communication, as its storage grid components may integrate or interface to relevant technologies.

Following the recent launch of the project, TextGrid is in an intense phase of design and specification. This process is meant to be transparent, and the participation of external initiatives is encouraged, thus tackling early its goal to establish a key infrastructure that links not only resources but also people.

Acknowledgments. The authors would like to thank the TextGrid team for their valuable thoughts and comments on a draft of this paper. TextGrid is partially funded by the German Federal Ministry of Education and Research (BMBF) under the D-Grid initiative by agreement 07TG01A-H. Responsibility for the contents of this publication rests with its authors.

References

[1] Gentzsch, W.: D-Grid, an E-Science Framework for German Scientists. Online publication: http://www.d-grid.de/ (January 2006)

[2] McCarthy, W.: What is humanities computing? Toward a definition of the field. (rev. 1998).

[3] Braungart, G., Eibl, K., Jannidis, F. (eds.): Jahrbuch für Computerphilologie 4. mentis Verlag Paderborn (2002)

[4] Fellbaum, C. (ed.): WordNet - An Electronic Lexical Database. (May 1998), ISBN 0-262-06197-X.

[5] Banks, T.: Web Services Resource Framework (WSRF). A Primer by the OASIS Technical Committee on WSRF. Draft 01 (2005)

[6] O'Reilly, T.: What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software. Online Publication: http://www.oreillynet.com/go/web2.0. (September 2005)

[7] Frommholz, I., Knezevic, P., Mehta, B., Niederée, C., Risse, T., Thiel, U.: Supporting Information Access in Next Generation Digital Library Architectures. In: Agosti, M., Schek, H.-J., Türker, C. (eds.): Digital Library Architectures: Peer-to-Peer, Grid, and Service-Orientation. Proceedings of the Sixth Thematic Workshop of the EU Network of Excellence DELOS (2004) pp. 49--60

[8] Foster, I., Kesselmann, C., Tuecke, S.: The Anatomy of the Grid Enabling Scalable Virtual Organizations. In: LNCS 2150 (2001).

[9] Richard Sietmann: Wissenschaft + Web-Services = e-Science. BMBF will Grid-Technologien außerhalb der Großforschung etablieren. In: c't, v 24, November 2005. pp 46-49.

[10] Grimm, C., Pattloch, M. (eds): Analyse von AA-Infrastrukturen in Grid-Middleware. Report of the D-Grid Integration Project, Focus Area 3.4 (March 2006).

[11] Morgan, R.L., Cantor, S., Carmody, S., Hoehn, W., Klingenstein, K.: Federated Security - The Shibboleth Approach. In: Educause 27(4) (2004)

[12] Paschoud, J.: Shibboleth and SAML: at last, a viable global standard for resource access management. In: New Review of Information Networking (January 2005)

[13] Barton, T., Basney, J., Freeman, T., Scavo, T., Siebenlist, F., Von Welch, Ananthakrishnan, R., Baker, B., Keahey, K.: Identity Federation and Attribute-based Authorization through the Globus Toolkit, Shibboleth, Gridshib, and MyProxy. In 5th Annual PKI R&D Workshop, April 2006.

[14] Foster, I.: Globus Toolkit Version 4: Software for Service-Oriented Systems. In: Proceedings of the IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779 (2005) pp. 2-13

[15] Polleres, A., Toma, I., Fensel, D.: *Modeling services for the Semantic Grid*. In: The Dagstuhl Seminar 05271 - Semantic Grid: The Convergence of Technologies (May 2005)

[16] Foster, I., Kesselman, C., Nick, J.M., Tuecke, S.: Grid services for distributed system integration. In: IEEE Computer 35(6) (June 2002)

[17] Clay Shirky's Internet Writings: Situated Software. First published on the "Networks, Economics, and Culture" mailing list (March 2004)

[18] Geldof, M.: The Semantic Grid: will Semantic Web and Grid go hand in hand?, Report for the unit "Grid Technologies" of the European Commission (2004)

[19] De Roure, D., Jennings, N. R. and Shadbolt, N. R.: The Semantic Grid: Past, Present and Future. Procedings of the IEEE 93(3) (2005) pp. 669-681