Grant agreement no. 211714

neuGRID

A GRID-BASED e-INFRASTRUCTURE FOR DATA ARCHIVING/ COMMUNICATION AND COMPUTATIONALLY INTENSIVE APPLICATIONS IN THE MEDICAL SCIENCES

Combination of Collaborative Project and Coordination and Support Action

Objective INFRA-2007-1.2.2 - Deployment of e-Infrastructures for scientific communities

Deliverable reference number and title: **D10.2 Gridified Toolbox Year 2 Portfolio and Report**

Due date of deliverable: **Month 22**

Actual submission date: **30th November 2009**

Start date of project: **February 1st 2008** Duration: **36 months**

Organisation name of lead contractor for this deliverable: **maat Gknowledge**

Revision: Version **1**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

Glossary

Term	Definition
IGT	Infrastructure Ground Truth. Level 0 of neuGRID's infrastructure hosting the GCC and DCC sites and associated test-bed
GCC	Grid Coordination Center. Core (common to all sites) services of the grid infrastructure
DCC	Data Coordination Center. Core (common to all sites) services of the database infrastructure
DACS	Data Archiving and Computational Site. neuGRID site offering and managering a set of physical resources
DCS	Data Collection Site. End-user sites acquiring data and connecting to a given DACS
Gridification	The engineering process of porting an existing application to the grid, so that it can be executed via the grid enactment environment
Pipeline	A pipeline is a set of data processing elements connected in series, so that the output of one element is the input of the next one (extracted from Wikipedia.org)
Workflow	A workflow is a depiction of a sequence of operations, declared as work of a person, work of a simple or complex mechanism, work of a group of persons, work of an organization of staff, or machines (extracted from Wikipedia.org)
Imaging Algorithm	An application, typically under the form of a Unix-like binary which manipulates imaging data
LORIS	The databasing software used in neuGRID, which offers interfaces to acquire and quality control data
Cortical Thickness	The thickness of the human cortex, estimated as the distance between the white matter surface and the grey matter surface across the entire cortex. This quantity can be used to analyse regional variations within and between subjects and/or populations.
Healthgrid	A grid-based environment in which data of medical interest can be stored and made easily available to different actors in healthcare systems such as physicians, healthcare centres, patients and citizens
SOA	Service-oriented Architecture
SOMA	Service-oriented Modelling and Architecture
VUmc	VU Medisch Centrum
KI	Karolinska Institute
FBF	Fatebenefratelli
PACS	Picture Archiving and Communication System
Voxel	A voxel (a combination of the words volumetric and pixel) is a volume element, representing a value on a regular grid in three dimensional space. This is analogous to a pixel, which represents 2D image data.
MINC	Medical Imaging NetCDF
NetCDF	Network Common Data Form
DICOM	Digital Imaging and COmmunications in Medicine
Analyse	Analyze is an image processing program and data format, written by The Biomedical Imaging Resource at the Mayo Foundation
GPL	General Public Licence
WSDL	Web Service Description Language
W3C	World Wide Web Consortium
Data Mining	Data mining is the process of extracting patterns from data. As more data are gathered, with the amount of data doubling every three years, data mining is becoming an increasingly important tool to transform these data into information. It is commonly used in a wide range of profiling practices, such as marketing, surveillance, fraud detection and scientific discovery.Definition extracted from: http://en.wikipedia.org/wiki/Data_mining

*Note: glossary terms are followed by a * symbol in the remainder of this document.*

Executive Summary

The present document attempts to list the gridified data mining* toolkits as they are ported to the grid by work package 10 (WP10) collaborators. By doing so, the intent is to establish neuGRID's portfolio of neuroimaging pipelines and thus to support scientists in better defining, planning and running new experiments in the infrastructure, while figuring out whether additional toolkits should be proposed for integration over year 3.

This first version of the document follows and addresses part of the relevant user and system requirements analysis (see section 2 for more details) while delivering the concrete list of available toolkits in the grid. It is anticipated that the document will go through a series of iterations in light of the ongoing gridification work and all along the neuGRID project lifetime, with a major version release at <u>Month 36</u>. The latter will constitute the neuGRID's portfolio of neuroimaging tools for post project exploitation.

In order to give clarity to this report, the remainder of the document is structured as follows:

1. Introduction
   - Outcome: deliverable context and positioning
2. Requirements Conclusions Recall
   - Outcome: recall of concerned requirements and rational to the presented work
3. Gridified Toolbox Portfolio
   - Outcome: list of available toolkits in the system
4. Conclusions and Future Work
   - Outcome: year 2 conclusions and up-coming portfolio extension

1. Introduction

This active document is meant to be refined all along the neuGRID project lifetime. It aims to provide the list of gridified data mining* toolkits and pipelines as well as associated documentation, which were identified as necessary tools for our end-users within the requirements specifications.

1.1. Purpose of the Document

This document lists the achieved gridification work over the two first years of the project execution. This work has been carried out within the tasks entitled "*T10.2 Algorithms Gridification*" and "*T10.3 Algorithm Pipeline Gridification*", which started at Month 7 and will complete at Month 36, with the following objective (extract from the project description of work):

T10.2: "*Develop the necessary 'glue' to tie together the gridification components, the distributed medical and brain image processing services, in a comprehensive, self-contained, secure and deployable package. Gridify and test the existing algorithms within the infrastructure, based on the gridification model defined in T10.1. Propose optimised/refined gridification models based on the test driven by WP11 in the infrastructure. Release the prototype software following the guidelines expressed in T11.1.The outcome of this task will be a gridified Toolbox, which will be refined/optimized until the end of the project (P4 MAAT, P2 NE, P3 UWE).*"

T10.3: "*Design and implement the necessary framework of services on top of the grid infrastructure, including meta-scheduler, high-availability and optimisation services adapted to the project requirements, capable of triggering algorithms executions across centres from the grid to local clusters. Implement the neuGRID Workflow Engine for managing the execution of different combinations of algorithms. This engine will interpret workflow process descriptions generated by end-users and will execute them in the grid infrastructure (P4 MAAT, P3 UWE).*"

In the following sections, the reader will gain understanding on the identified and now available toolkits of neuroimaging algorithms and pipelines that neuGRID offers to end-users. It establishes a classification, and as such, it is intended to be a useful reference throughout the development and test of new pipelines by neuroscientists.

1.2. Document Positioning and Intended Audience

WP10 "*Algorithms and Pipeline Gridification*" aims to make existing brain image analysis algorithms compatible with the grid environment (the so-called process of "gridification") and to develop the necessary foundations for their publication, management, sharing, combination, and scheduling in the neuGRID system. The outcome of this work package will be an algorithm "*toolbox*" made available to end-user communities through the platform, which can be discovered, enriched, invoked and applied in different ways onto a large set of images and clinical records.

More precisely, the work package aims to (extract from the project description of work) *(1) evaluate the existing algorithms' implementations and requirements in terms of software adaptation and interfacing, (2) design and implement a set of distributed and cooperative optimization methods for facilitating algorithms gridification and their future scheduling within the platform, (3) design and implement a set of interfaces for managing the algorithms in the grid (from algorithm publication, to versioning, to training, to sharing), (4) gridify the algorithms, re-engineer algorithms inner interfaces so that their input/output can be plugged in the grid, (5) define adapted scheduling policies for the selected algorithms, based on algorithms requirements, (6) design and develop a grid-based workflow management system for combining, optimizing,* *executing and monitoring algorithms, which extends the workflow capabilities of the grid middleware and addresses fully the brain image processing algorithms needs, and last but not least (7) extend the scheduling possibilities.*

Thus, the presented report is destined to serve in priority all protagonists of the Joint Research (JRA) and Services (SA) activities of the project, more particularly to IT researchers, IT developers and neuroscientific algorithm developers involved in the following work packages:

Services Activities – SA

WP Id	*WP Title*	*WP10 Contribution*
WP5	Brain Imaging Services Provision	To provide a gridification model specification and corresponding implementation for brain imaging services to be published, discovered and executed in neuGRID's platform
WP6	Distributed Medical Services Provision	To provide a Workflow/ Pipelining service and corresponding API for higher level distributed medical services to submit, execute and monitor algorithms/ pipelines
WP7	Grid Services Provision	To dictate the deployment of necessary underlying grid services and corresponding configurations
WP8	Deployment Services Provision	To dictate the deployment of necessary underlying neuGRID services and corresponding configurations

Joint Research Activities – JRA

WP Id	*WP Title*	*WP Relation*
WP9	User and System Requirements Analysis	To conform with requirements analysis conclusions
WP11	Platform Integration, Performance and Feasibility Tests	To support neuGRID services integration

To a lesser extent, since indirectly concerned (through the natural abstraction of Workflow/ Pipeline authoring environments such as the ones proposed in WP6, see D6.1 for more details), the neuroscientists and prospective users (e.g. Pharmaceutical industries) as well as inside and outside reviewers of the project activities, are anticipated as potential additional audience.

1.3. Reference Documents

Prior to reading this document the reader should be familiar with additional documents/ deliverables produced within the neuGRID project, which have or are considered to potentially impact on the design and developments of WP10. The following is a list of such documents sorted by information sources, activities and corresponding work packages (Note: list of available documents at the time of writing):

Services Activities Related Documents

WP Id	*WP Title*	*Documents*
WP5	Brain Imaging Services Provision	D5.1. Brain Imaging Service Portfolio Specification Document
WP6	Distributed Medical Services Provision	D6.1. Design Document including API Documentation and Description of Functionality for the Underlying Layer
WP7	Grid Services Provision	D7.1. Test-bed Installation and API Documentation
WP8	Deployment Services Provision	D8.1. Ground Truth and Phase 1 Deployment Test and Validation Report

Joint Research Activities Related Documents

WP Id	*WP Title*	*Documents*
WP11	Platform Integration, Performance and Feasibility Tests	D11.1. AC/DC1 and Story Lines Test Suite Specification and Report neuGRID Architectural Considerations Presentation (see neuGRID CMS, WP11 directory: https://www.neugrid.eu/owl-0.90)

Other Related Documents

Title	*Documents*
Project Documents	Project Description of Work
Requirements Analysis Supporting Material	Workflow related requirements analysis material, including recordings, meeting minutes and pipelines examples (see neuGRID CMS, WP10 directory: https://www.neugrid.eu/owl-0.90)

2. Extract from the Requirements Conclusions

This section aims to recall former requirements analysis conclusions as described in deliverable D10.1 "*Gridification Model Specification*" at Month 14. In particular, it focuses on the identified toolkits and pipelines being used by the 3 involved clinical research centers.

2.1. Identified Toolkits and Pipelines

The following table lists the pipeline tools that are in frequent use by the research centers as expressed by end-users. It aims to give a taste on the faced difficulty and heterogeneity of available imaging/ mining toolkits, whether being commercial suites or community software:

Institute	Pipeline Tools	Analysis Tools
VUmc	fMRIB Software Library (FSL): Flirt, Fnirt, FDT, FAST, Melodic (visualization tool), Siena, XSiena, FEAT, http://www.fmrib.ox.ac.uk • MRIcro, Brain Extraction Tool (BET), http://www.sph.sc.edu/comd/rorden/mricro.html • Montreal Neurological Institute (MNI) (BIC Tools & Software – The Brain Imaging Software Toolbox): N3. http://www.bic.mni.mcgill.ca/software/ • BioInformatics Research Network (BIRN) (Gradiant Non-Linearity Distortion Correction): Gradient non-linearity. http://www.nbirn.net/ • DRG Fluid. • Generic: ○ Image calculations (adding subtracting, multiplying etc) ○ Morphological operations on images ○ File format conversions	Statistical Parametric Mapping – SPM http://www.fil.ion.ucl.ac.uk/spm/software/
KI	• MNI BIC Tool – CIVET Pipeline http://wiki.bic.mni.mcgill.ca/index.php/CIVET , • FSL, • Brainvoyager http://www.brainvoyager.com/ • Matlab http://www.matlab.com , • Analysis fo Functional NeuroImages (AFNI), http://afni.nimh.nih.gov/afni/ • E-prime http://www.pstnet.com/ and • Statistica.	Hermes (Hermes Medical) B-MAP (Pipeline 1 and Pipeline 2) http://www.hermesmedical.com/

FBF	• FSL Tools fMRIB's Diffusion Toolbox FDT 2.0, Melodic
	
	• MNI BIC Tools:
	  • Display, register, Brainsuite
	
	• LoNI http://www.loni.ucla.edu/Software/ tools:
	  • Dual_warpe_warpcurve, Decoder_blend_all, mk_seg16bit, mk_gray, add_gray_to_inflated_LEFT1, add_gray_to_inflated_RIGHT1, pmap_apeVSctrl, make_UVL_*; 1st_script_tracer_avg_DIAG; 2nd_script_core_test_L_DIAG; 2nd_script_core_test_R_DIAG; Pmap_DistCore_DIAG
	
	• MRIcro (MRIcro) (visualization)
	  • BET Function
	
	• IdeALab Tools (IdeALab) http://neuroscience.ucdavis.edu/idealab/software/index.php
	
	• Image Conversion software
	  • MRIconverter
	  • dcm2nii
	
	• New Promising Tools:
	  • 3D Slicer, VTK, Freesurfer, MPIAV, NA-MIC Kit components, MED-INRIA, BrainVoyager, BrainMAP
		• SPSS http://www.spss.com/,
		
		• Statistical Parametric Mapping – SPM, Matlab, Quanta 6.1
		
		• R (R) http://www.r-project.org
		
		• Statistical Parametric Mapping – SPM

This list demonstrates that end-users develop preferences over time from personal experience and projects, which lead them to use various combinations of toolkits/ algorithms to extract complex features. From these only three centers short survey, one can notice that there are however a few common tools, as highlighted in the following table.

	FSL	MNI/CIVET	LoNI	SPM	MRIcro/BET	SPSS	HERMES	Idealab	Matlab	R	AFNI	E-Prime	Statistica	DRG	BIRN	BrainVoy.	FreeSurfer
VUmc	X	X		X	X									X	X		
KI	X	X					X		X		X	X	X			X	
FBF	X	X	X	X	X	X		X	X	X							X

As a conclusion to this initial comparative table, FSL, MNI/BIC, SPM, MRIcro and Matlab seem to be the most common set of (neuro) imaging and data mining* toolkits being used by our neuroscientists. It also demonstrates that gridifying a single toolkit would not be of any relevance to our end-users. Indeed, neuroscientists are used to mix toolkits and algorithms in their daily research experiments. This is a reality that WP10 is addressing by gradually porting the mostly used toolkits to the grid. <u>This list has therefore been used to prioritize gridifications, although WP10 remains at end-users' disposal and thus have been and will still be integrating on-demand and as project resources allow for this.</u>

2.2. Categorization of Toolkits and Pipelines

The following table gives a list of popular toolkits and corresponding image processing capabilities used to respectively normalize data, convert image files, anonymize data, extract features from within images and process statistics. This classification aims to introduce the notion of categories that the resulting neuGRID system could use to structure the gridified algorithms portfolio.

Main Category	Type of Processing	Pipeline / Algorithm	Toolkit
Pre & Intermediary Processing	Normalization	Linear and nonlinear (correction factors)	SPM
		Segmentation (voxels labelling priors-based)	SPM
		Warping (sulci based)	LoNI
		Warping (intensity based)	MNI
	File Conversion	Dicom to MINC	MNI
		Dicom to Analyze	MNI
			MRIcro
	Anonymization	Face Scrambling	MNI, FreeSurfer
		Pseudonymization	LoNI
Feature Extraction	Segmentation	Cortical Density	SPM
			LoNI
		Cortical Thickness	LoNI
		Hippocampus Atrophy (shrinkage)	LoNI
		Hippocampus Volume	MNI
			LoNI
		White Matter Volume and Distribution	IdeALab
		Cortical Thickness	MNI
		Cortical Contour Drawing + Voxels Counting	MNI
		White Matter Age Related Scale (Wahlund)	--
		Regional Brain Metabolism Alterations	HERMES
Post Processing and Analysis	Statistics	Cross Population Patterns	R

The next section describes the present portfolio of neuroimaging toolkits as available in neuGRID's infrastructure which can be discovered, invoked and executed in the grid. The portfolio is then presented in a structured manner using the above classifications.

3. Gridified Toolbox Portfolio

The following sections list the different toolkits so far gridifed and thus available within the neuGRID infrastructure for execution. The toolkits are classified according to their nature and more technical information is provided.

3.1. Neuroimaging Toolkits

3.1.1. McGILL/MNI – The CIVET Toolkit

McGILL/MNI – CIVET
The CIVET project was initiated at the McConnell Brain-Imaging Centre, Montreal Neurological Institute, McGill University (BIC/MNI) as an image processing "pipeline" focused on the extraction of the cortical thickness in the human brain from MRI data. Having grown into a more general structural MRI processing system, CIVET is now an environment that provides easy access to many of the BIC/MNI software. The objective of CIVET was to make it possible for someone with little or no programming background to make full use of the available software for automated structural (anatomical) research, while simultaneously allowing developers to have maximal capacity to customize, add or improve various functions to the platform.
More information available at: http://wiki.bic.mni.mcgill.ca/index.php/CIVET
Category (based on section 2.2): Feature Extraction / Segmentation
Key Pipeline Example: CIVET Cortical Thickness Extraction
Location within the grid Worker Nodes: **/opt/Quarantines/200906/bin** Architecture: both 32-bit and 64-bit
Comments: The CIVET pipeline for cortical thickness extraction has been provided with a wrapping shell script, which takes care of optimizing the pipeline execution over large datasets. The script encompasses the logic to check whether the pipeline has already been applied to given datasets within the grid. The script is located at **/bin/civet-launch.sh** and accepts the following parameters: *1. LFC host,* *2. LFN of input data (/grid/neugrid/..../xxxxx.mnc.gz file),* *3. LFN directory where to store ouput data (/grid/neugrid/..../ directory),* *4. civet prefix to use.*

3.1.2. MGH/MCBI – The FreeSurfer Toolkit

MGH/MCBI - FreeSurfer
FreeSurfer is a set of software tools for the study of cortical and subcortical anatomy. In the cortical surface stream, the tools construct models of the boundary between white matter and cortical grey matter as well as the pial surface. Once these surfaces are known, an array of anatomical measures becomes possible, including: cortical thickness, surface area, curvature, and surface normal at each point on the cortex. The surfaces can be inflated and/or flattened for improved visualization. The surfaces can also be used to constrain the solutions to inverse optical, EEG and MEG problems. In addition, a cortical surface-based atlas has been defined based on average folding patterns mapped to a sphere. Surfaces from individuals can be aligned with this atlas with a high-dimensional nonlinear registration algorithm. The registration is based on aligning the cortical folding patterns and so directly aligns the anatomy instead of image intensities. The spherical atlas naturally forms a coordinate system in which point-to-point correspondence between subjects can be achieved. This coordinate system can then be used to create group maps (similar to how Talairach space is used for volumetric measurements). Most of the FreeSurfer pipeline is automated, which makes it ideal for use on large data sets.
More information available at: http://surfer.nmr.mgh.harvard.edu/
Category (based on section 2.2): Feature Extraction / Segmentation
Key Algorithm Example: DEFACE Face Scrambling
Location within the grid Worker Nodes: **/opt/freesurfer/bin** Architecture: both 32-bit and 64-bit
Comments: the FreeSurfer toolkit, beyond being a widely used software package in the neuroimaging community, was primarily installed in order to prototype the face scrambling service which will be used to anonymize neuGRID's data.

3.2. Statistical Toolkits

3.2.1. GNU/FSF – The R Project Toolkit

GNU/FSF - R
R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity. One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.
More information available at: http://www.r-project.org/
Category (based on section 2.2): Post Processing & Analysis / Statistics
Key Pipeline Example: not applicable
Location within the grid Worker Nodes: **/usr/bin/R/bin** Architecture: both 32-bit and 64-bit
Comments: the R Project statistical toolkit is provided as it is the toolkit of choice for performing statistical analyses on results obtained with the CIVET pipeline.

3.3. Gridified Toolkits vs Requirements

The following table recalls the toolkits identified during the requirements analysis, as is referenced in former section 3.1. The red color is used to differentiate gridified toolkits from the rest.

	FSL	MNI/CIVET	LoNI	SPM	MRIcro/BET	SPSS	HERMES	Idealab	Matlab	R	AFNI	E-Prime	Statistica	DRG	BIRN	BrainVoy.	FreeSurfer
VUmc	X	X		X	X									X	X		
KI	X	X					X		X		X	X	X			X	
FBF	X	X	X	X	X	X		X	X	X							X

As can be seen from above, one of the most important pipeline toolkit (i.e. CIVET) has been gridified, as well as two lower priority toolkits (i.e. R Project and FreeSurfer).

Indeed, the priorities have been reconsidered in light of new requirements expressed by end-users at FBF over year 2 and according to neuGRID's developments needs, in particular concerning the anonymization (face scrambling) prototyping efforts.

4. Conclusions and Future Work

In conclusion, no major issues were faced in gridifying applications so far. Problems were rather encountered at execution time in particular with library dependencies at the Worker Nodes' OS level. This is the case for CIVET in 64-bit version which still is being debugged following the AC/DC2 data challenge execution, where a 10% failure rate was reached from over 6'300 executions. Whereas part of these failures were definitely due to bad quality data, library incompatibilities have played a role in some of the cases.

Deliverable D10.3 "*Gridified Toolbox Year 3 Portfolio and Report*", planned at Month 36 will report on the findings and conclusions of our current investigation of the CIVET execution problem.

In terms of future work and at the time of writing, work is already ongoing in the gridification of additional toolkits such as ITK/VTK, MatLab and FSL. The objective is to port all these applications over the coming months and to test them through the AC/DC3 challenge together with WP11 collaborators. Similarly to what was done over year 2, new toolkits to be gridified will be prioritized according to both the requirements specifications and end-users new demands.

Bibliographical References

[1] Arsanjani, A. (2004). Service-oriented modeling and architecture, IBM technical article. Retrieved October 20, 2008, from http://www.ibm.com/developerworks/library/ws-soa-design1/ .

[2] C. M. MacKenzie et al. Reference Model for Service Oriented Architecture 1.0, OASIS Committee Specification 1, 2 August 2006

[3] Papazoglou, M.P. (2003). Service-oriented computing: concepts, characteristics and directions. In *Proceedings of the 4th International Conference on Web Information Systems Engineering, WISE 2003, Roma, Italy*.

[4] McClatchey, R., Manset, D., & Solomonides, T. (2006). Lessons learned from MammoGrid for integrated biomedical solutions. In *19th IEEE Int. Symposium on Computer-Based Medical Systems, CBMS 2006*.

[5] Health-e-Child, The EU FP6 Information Societies Technology Project. (2008). Retrieved October 20, 2008, from http://www.health-e-child.org/ .

[6] "Health-e-Child: A Grid Platform for European Paediatrics" K. Skaburskas, F. Estrella, J. Shade, D. Manset, J. Revillard, A. Rios, A. Anjum, A. Branson, P. Bloodsworth, T. Hauer, R. McClatchey, D. Rogulin Proceedings of the 2007 Computing in High Energy and Nuclear Physics International Conference - CHEP07, (USA), 2007 Publication in Journal of Physics: Conference Series.

[7] MRIcro. http://www.sph.sc.edu/comd/rorden/mricro.html

[8] FMRIB Software Library. http://www.fmrib.ox.ac.uk

[9] http://people.cs.uchicago.edu/~hai/tmp/doc/siena/index.html

[10] Oliveira, M. C., Cirne, W., & de Azevedo Marques, P. M. (2007). Towards applying content-based image retrieval in the clinical routine. *Future Gener. Comput. Syst.*, 23 (3), 466-474.

[11] Bellotti, R., Cerello, P., Tangaro, S., Bevilacqua, V., Castellano, M., Mastronardi, G., De Carlo, F., Bagnasco, S., Bottigli, U., Cataldo, R., Catanzariti, E., Cheran, S. C., Delogu, P., De Mitri, I., De Nunzio, G., Fantacci, M. E., Fauci, F., Gargano, G., Golosio, B., Indovina, P. L., Lauria, A., Lopez Torres, E., Magro, R., Masala, G. L., Massafra, R., Oliva, P., Preite Martinez, A., Quarta, M., Raso, G., Retico, A., Sitta, M., Stumbo, S., Tata, A., Squarcia, S., Schenone, A., Molinari, E., & Canesi, B. (2007). Distributed medical images analysis on a Grid infrastructure. *Future Gener. Comput. Syst.*, 23 (3), 475-484.

[12] Budura, A., Cudré-Mauroux, P., & Aberer, K. (2007). From bioinformatic web portals to semantically integrated Data Grid networks. *Future Gener. Comput. Syst.*, 23 (3), 485-496.

[13] glite, A lightweight middleware for grid computing. (2008). Retrieved October 20, 2008, from http://glite.web.cern.ch/glite/ .

[14] The Globus Alliance: http://www.globus.org/

[15] "Gridifying Biomedical Applications: Experiences of the Health-e-Child Project" D. Manset, F. Pourraz, A. Tsymbal, J. Revillard, K. Skaburskas, R. McClatchey, A. Anjum, A. Rios, M. Huber Handbook of Research on Computational Grid Technologies for Life Sciences, Biomedicine and Healthcare. Information Science Reference (IGI Global). Submitted - still being evaluated.

[16] Tristan Glatard, Johan Montagnat, Xavier Pennec, David Emsellem, Diane Lingrand. *"MOTEUR: a data-intensive service-based workflow manager"* Research Report I3S, number I3S/RR-2006-07-FR, 37 pages, Sophia Antipolis, France, mar 2006

[17] Diane Lingrand, Johan Montagnat, Tristan Glatard. *"Estimating the execution context for refining submission strategies on production grids"* (workshop) in Proceedings of the Assessing Models of Networks and Distributed Computing Platforms (ASSESS / ModernBio) (CCgrid'08), pages 753 -- 758

[18] Tristan Glatard, David Emsellem, Johan Montagnat. *"Generic web service wrapper for efficient embedding of legacy codes in service-based workflows"* (workshop) in Proceedings of the Grid-Enabling Legacy Applications and Supporting End Users Workshop (GELA'06), pages 44--53, Paris, France, jun 2006

[19] LoNI Software Tools. http://www.loni.ucla.edu/Software/

[20] "A Formal Model-Driven Approach for Grid Application Architectures" D. Manset, H. Verjus. Technical Report LISTIC No 07/02, University of Savoie - LISTIC, 2007.

[21] "Managing Separation of Concerns in Grid Applications Through Architectural Model Transformations" D. Manset, H. Verjus & R. McClatchey. Lecture Notes in Computer Science Vol 4758 pp 308-310 ISBN ISBN 978-3-540-75131-1 Springer-Verlag, 2007.