



GISELA

1ST YEAR ACTIVITY HIGHLIGHTS ON VRCs AND USER SUPPORT

EU DELIVERABLE: D3.2

Document Full name	GISELA-D3.2-v1.8
Date	26/08/2011
Activity	WP3 / User Communities Support
Lead Partner	UFRJ
Document status	APPROVED
Classification Attribute	PU (Public)
Document link	http://documents.gisela-grid.eu

Abstract: This document reports on the WP3 achievements from M1 to M12 (September 2010 - August 2011), describing the results on promoting knowledge dissemination and supporting VRCs.



Copyright notice

Copyright © Members of the **GISELA** Consortium, 2010

GISELA (“Grid Initiatives for e-Science virtual communities in Europe and Latin America”) is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. **GISELA** began on 1st September 2010 and will run for 2 years.

For more information on GISELA, its partners and contributors please see www.gisela-grid.eu.

You are permitted to copy and distribute, for non-profit purposes, verbatim copies of this document containing this copyright notice. This includes the right to copy this document in whole or in part, but without modification, into other documents if you attach the following reference to the copied elements: “Copyright © Members of the **GISELA** Consortium, 2010. See www.gisela-grid.eu for details”.

Using this document, in a way and/or for purposes not foreseen in the paragraph above, requires the prior written permission of the copyright holders.

The information contained in this document represents the views of the copyright holders as of the date such views were published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS “AS IS” AND ANY EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE MEMBERS OF THE **GISELA** COLLABORATION, INCLUDING THE COPYRIGHT HOLDERS, OR THE EUROPEAN COMMISSION BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THE INFORMATION CONTAINED IN THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Delivery Slip

	Name	Partner/Activity	Date	Signature
From	WP3	UFRJ / WP3 – User Communities Support		
Reviewed by	Technical Board			
Approved by	Management Board		26/08/2011	B. Marechal Ph. Gavillet S. Jalife Villalón L. A. Trejo Rodriguez R. Barbera R. Ramos Pollán

Document Log

Issue	Date	Comment	Author
0-1	28/06/2011	First draft	B. Marechal
0-2	23/07/2011	Second draft, content distribution	Leandro Ciuffo
0-3	30/07/2011	Third draft, including contributions from Rafael Mayo e Guillermo Diaz	Leandro Ciuffo
0-4	04/08/2011	Tables of sections 3 and section 4 inserted	Diego Scardaci
0-5	05/08/2011	Revision of section 3.4, including latest contributions from Rafael Mayo. Executive summary added	Guillermo Diaz Herrero
0-6	07/08/2011	Overall revision, figures updated, conclusion.	Leandro Ciuffo
0-7	13/08/2011	Incorporation of contributions from Jesus Cruz, Philippe Gavillet, Ramon Diacovo and Harold Castro.	Leandro Ciuffo
1-7	23/08/2011	Final review and approval	B. Marechal

Document Change Record

Issue	Item	Reason for Change

TABLE OF CONTENTS

1. INTRODUCTION	6
1.1. PURPOSE OF THE DOCUMENT	6
1.2. DOCUMENT ORGANISATION	6
1.3. APPLICATION AREA	6
1.4. DOCUMENT AMENDMENT PROCEDURE	6
1.5. GLOSSARY	7
2. EXECUTIVE SUMMARY	8
3. ACHIEVEMENTS	9
3.1. VIRTUAL RESEARCH COMMUNITIES SUPPORTED	9
3.1.1. <i>Life-Sciences VRC</i>	10
3.1.2. <i>Earth-Sciences VRC</i>	10
3.1.3. <i>HEP VRC</i>	11
3.2. OVERALL LIST OF APPLICATIONS AND VRCS	12
3.2.1. <i>Usage statistics</i>	13
3.3. APPLICATIONS PORTFOLIO	14
3.4. USER SUPPORT	19
3.4.1. <i>Documentation</i>	19
3.4.2. <i>Publications</i>	21
3.4.3. <i>Training events</i>	21
4. SCIENCE GATEWAY FOR INDUSTRIAL APPLICATIONS	23
5. HUMAN EFFORT	24
6. OPEN ISSUES AND / OR DEVIATIONS FROM THE WORK PLAN	27
7. PLANS FOR THE NEXT REPORTING PERIOD	28
7.1. DEVELOPMENTS AT UNIANDES (COLOMBIA)	28
8. CONCLUSIONS	30

TABLE OF FIGURES

Figure 1: Number of supported VRCS	9
Figure 2: Number of application groups from Latin America	13
Figure 3: Cumulative CPU-hours (normalized) grouped by VOs	14
Figure 4: Monthly utilization by each VO (percentage)	14
Figure 5: The EGI Applications Database (see http://appdb.egi.e/)	15
Figure 6: Project Training Material search portal	19
Figure 7: Number of self-training material made available through Documents Server	20
Figure 8: Number of Training Events	22
Figure 9: Reference model of the Science Gateway for the Industrial Applications	23



TABLE OF TABLES

Table 1: WP3 first year quality metrics.....	8
Table 2: GISELA sites supporting HEP VOs (list obtained by querying the GISELA BDII on May 24th 2011)	11
Table 3: VRCS' Applications.....	12
Table 4: GISELA Applications belonging to status "S5".....	15
Table 5: GISELA Applications belonging to status "S4".....	18
Table 6: List of training events organised in the first year.....	21
Table 7: Human effort allocated to WP3 management	24
Table 8: List of application's developers/supporters.....	24

1. INTRODUCTION

1.1. PURPOSE OF THE DOCUMENT

This document aims at reporting the achievements, the current status of the user support activities, as well as the work carried out by WP3 as a whole over the period from 01/09/2010 (M01) to 31/08/2011 (M12).

For a comprehensive view of the Project and of the GISELA Consortium, the Description of Work (DoW)¹ and the Consortium Agreement (CoA)² should be consulted.

1.2. DOCUMENT ORGANISATION

Section 2 gives an Executive Summary of the document. Section 3 presents the achievements of the work carried out by WP3. Section 4 describes the development of a pilot Science Gateway. In section 5 the WP3 manpower available for WP3 is presented. Section 6 discusses the deviations from the initial work plan. Section 7 introduces the plans for the next reporting period. Finally, Section 8 presents some conclusions.

1.3. APPLICATION AREA

The target audience for this document is:

- The members of the Project;
- The European Commission Services;
- The Project Reviewers;
- The External Advisory Committee (EAC);
- The general public.

1.4. DOCUMENT AMENDMENT PROCEDURE

Amendments to this document can be requested by any Project Member to the Project Coordinator, via the Project Office (hlp-gisela@hlpdeveloppement.fr).

¹ <http://documents.gisela-grid.eu/record/32?ln=en>

² Consortium Agreement (CoA) available upon request to the GISELA Project Office (hlp-gisela@hlpdeveloppement.fr)

1.5. GLOSSARY

CoA	Consortium Agreement
DoW	Description of Work
EAC	External Advisory Committee
CLARA TT	RedCLARA's Transition Team
GOC	Grid Operational Centre
GSC	Grid Support Centre
JSR	Java Specification Request
HEP	High Energy Physics
INPA	The Brazilian institute for the Research of the Amazon-region
LGI	Latin American Grid Initiative
NGI	National Grid Initiatives
PC	Project Coordinator
PO	Project Office
RC	Resource Centre
ROC	Regional Operation Centre
TB	Technical Board
TC	Technical Coordinator
UFRJ	Universidade Federal do Rio de Janeiro - Brasil
VO	Virtual Organisation
VRC	Virtual Research Community
WP	Work package
WP1	Administrative and Technical management of the Project
WP2	Dissemination and Outreach
WP3	User Communities Support
WP4	NGI / LGI Infrastructure Services
WP5	Network Resource Provision
WP6	Infrastructure and Applications-oriented Services for User Communities

2. EXECUTIVE SUMMARY

The GISELA infrastructure currently supports 3 VRCS (Life Science, Earth Science and HEP), which were primary prospected to address the demand of Latin American research groups. Those VRCS serves a set of 19 applications groups from Latin America.

Most of the legacy applications from the former EELA-2 project have been registered in the new world-wide applications databases provided by the European Grid Initiative (EGI). Such a database works as an “application directory”, available for any GISELA user interested in trying one of the listed applications.

As an attempt to increase Grid adoption and usage, a pilot Science Gateway for the Industrial Application “Industry@Grid” has been developed in collaboration with WP6.

With respect to the accomplishment of foreseen quality metrics, WP3 is in good shape at the end of first year, despite of the difficulties reported during the Crash Program meeting³.

Table 1: WP3 first year quality metrics

Quality metric	Current status Y1Q4	Expected outcome Y1	Expected outcome Y2
1 Number of supported VRCS	3	≥ 2	≥ 3
2 Number of application’s groups from Latin America	19	≥ 15	≥ 36
3 Number of training events	3	≥ 1	≥ 4
4 Number of self-training material	40	≥ 20	≥ 40
5 Number of papers published (WP3 related)	5	≥ 5	≥ 20

Signed MoUs with EPIKH^{4,5} and EGI-InSPIRE^{6,7} projects, as well as close cooperation with the CHAIN⁸ project, much helped in the accomplishment of first year quality metrics. However, more intense involvement from the CLARA TT will be crucial to (1) increase number of new application’s groups from LA, (2) foster a high and stable demand of GISELA pledged resources, and (3) take the baton on providing specialized Support Services before the project ends.

Finally, plans for the second year are proposed seeking the success of WP3 activities in contributing to the long-term sustainability of e-Infrastructures in Latin America.

³ “Crash Program Final Report. Follow up at the end of Q3” at <http://documents.gisela-grid.eu/record/218>

⁴ <http://documents.gisela-grid.eu/record/166>

⁵ <http://www.epikh.eu>

⁶ <http://documents.gisela-grid.eu/record/171>

⁷ <http://www.egi.eu/projects/egi-inspire/>

⁸ <http://www.chain-project.eu>

3. ACHIEVEMENTS

3.1. VIRTUAL RESEARCH COMMUNITIES SUPPORTED

Following the EELA-2 track, GISELA put its focus on supporting VRCs, in particular those of interest from Latin American researchers.

The WP3 activities started by prospecting known VRCs to become or continue to be a GISELA “customer”. Direct contacts with already well established VRCs (such as We-NMR or HealthGrid) or group of researchers were carried out.

At the same time, a first draft of a list of contacts (researchers) was elaborated by the project and, after consensus, those people were approached in order to know their availability to join an existing VRC or even to form a (small) one, from individual institutions. From the original list, only two research groups replied positively, both from Brazil: astrophysicists from UFRJ and climate researchers from INPA.

As reported in D3.1, three VRCs were early identified to be supported:

- High Energy Physics (VOs: Auger, ALICE; ATLAS; CMS, LHCb)
- Life Sciences (VO: WeNMR and Biomed)
- Earth Sciences (VO: supported via GISELA's catch-all virtual organisation - prod.vo.eu-eela.eu).

Figure 1 compares the number of Virtual Research Communities supported by GISELA to what has been committed in the DoW.

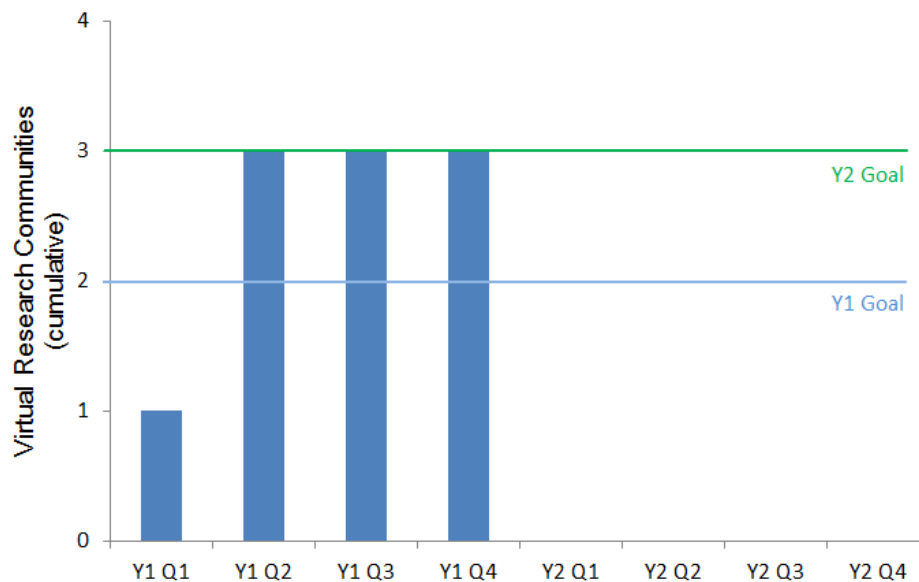


Figure 1: Number of supported VRCs

A direct link with the CHAIN project has been established. One of its objectives is the consolidation of VRC support, not only in Latin America, but also in Asian and Africa. This way, continuous exchange of information has been provided and all the steps taken in the CHAIN side have been reported to GISELA in order to profit for the future and vice versa. CHAIN has carried out a big survey around the world, the results of which are expected for the next weeks. In this questionnaire, a specific section for VRC was present for identifying needs and requirements of different big users, NGIs and ROCs.

Lately, a workshop for putting together what VRCs need and what the Infrastructure providers can offer was organised by GISELA and CHAIN during the EGI User Forum 2011 (<https://www.egi.eu/indico/materialDisplay.py?contribId=183&sessionId=19&materialId=slides&confId=207>).

Further information about each supported VRC is provided in the next sections.

3.1.1. Life-Sciences VRC

The outcomes of the collaboration started during the EGI Technical Forum with WeNMR (<http://www.wenmr.eu/>: A worldwide e-Infrastructure for NMR and structural biology) was twofold: it enabled GISELA users to access WeNMR application portals and opened LA sites to the WeNMR VO (called enmr.eu VO), providing computing power to the collaboration. GISELA and WeNMR identified a NMR team in Brazil that showed interest in using the WeNMR portals. GISELA started helping it with digital identity management, which enabled its access to the WeNMR portal.

The UFRJ-IF site of the GISELA infrastructure supports this VO now, turning this site as the first to support the enmr.eu VO in Latin America.

An agreement with Healthgrid⁹ was also concluded. It comprises the following actions:

- Extend the support of biomed VOs to other sites in LA;
- Receive support from Healthgrid to provide Workflow tools to LA users;
- Exchange of services and software between both communities.

3.1.2. Earth-Sciences VRC

The Weather Research and Forecasting (WRF¹⁰) modelling system was selected to be the first application from the Earth Science Community. The WRF system is in the public domain i.e. freely available for community use. It is designed to be a flexible, state-of-the-art atmospheric simulation system. It is suitable for use in a broad range of forecasting applications ranging from meters to thousands of kilometres. It includes simulations for regional and global applications, parameterization, data assimilation, forecast and hurricane research, coupled-model applications

GISELA contacted 14 research groups in Latin America in order to understand their interest in using the WRF4G¹¹ (grid adapted version of WRF). Only the INPA institute declared interested in using WRF4G to perform simulations applied to the Amazon region. WP3 team is currently supporting them to deploy and run the first simulations. Further details are provided in section 7.

Besides this initiative, GISELA is working to leverage on Earth Science communities formed during EELA and EELA-2. Cuban groups (see <http://www.isgtw.org/feature/cuba-could-breathe-cleaner-air-thanks-grid>) are using the GISELA infrastructure to run WRF4G and AERMOD applications in order to improve the understanding on how the pollution affects the air quality in the atmosphere.

⁹ <http://www.healthgrid.org>

¹⁰ <http://www.mmm.ucar.edu/wrf/users/>

¹¹ <http://www.meteo.unican.es/software/wrf4g>

The WP3 manager has also attended the event *Role of e-Infrastructures for Climate Change Research* (Trieste Italy 16-20 May), presenting GISELA to more than 120 people (<http://documents.gisela-grid.eu/record/224>).

3.1.3. HEP VRC

GISELA supports each of the four LHC experiments VOs and the Pierre Auger application, a tool developed to manage the Grid platform devoted to this observatory located in Malargüe, Argentina, for studying ultra-high energy cosmic rays.

Table 2 gives the GISELA sites supporting HEP Virtual Organisations.

Table 2: GISELA sites supporting HEP VOs
(list obtained by querying the GISELA BDII on May 24th 2011)

HEP Virtual Organisation	GISELA CEs
alice	ce01-tic.ciemat.es cream-ce.ct.infn.it grid012.ct.infn.it tochtli.nucleares.unam.mx tochtli64.nucleares.unam.mx
atlas	ce.labmc.inf.utfsm.cl
auger	cale.uniandes.edu.co ce01.eela.if.ufrj.br ce02.eela.if.ufrj.br grid012.ct.infn.it tochtli.nucleares.unam.mx tochtli64.nucleares.unam.mx
cms	cale.uniandes.edu.co grid012.ct.infn.it
lhcb	ce01-tic.ciemat.es ce01.eela.if.ufrj.br grid012.ct.infn.it

3.2. OVERALL LIST OF APPLICATIONS AND VRCS

Table 3 presents the applications associated to each supported VRCS.

Table 3: VRCS' Applications

VRCS	Legacy applications	Virtual Organisations
Life Sciences	Haddock, Cyana, Xplor-NIH, CS-ROSETTA, MD, Amber, MDD, Bowtie, etc.	biomed, enmr.eu
Earth Sciences	CAM and WRF	prod.vo.eu-eela.eu
HEP	Applications managed by each VO	ALICE, ATLAS, AUGER, CMS and LHCb
Catch-all	CLARA Communities, Mexican Industry Community, Applications inherited from EELA and EELA-2, etc.	prod.vo.eu-eela.eu

Figure 2 shows the evolution of supported groups in Latin America. The figure also introduces two levels of support: the central one and the local one. The former regards the support offered by the WP3 task leaders while the latter refer to the support of local experts inside the institution that develops the application. Such a distinction of support levels is foreseen in the Latin American Grid Initiative (LGI) sustainability model¹²: local support being at RC or NGI/GOC levels; central support being at the LGI/GSC level.

¹² <http://documents.eu-eela.eu/record/1333/files/>

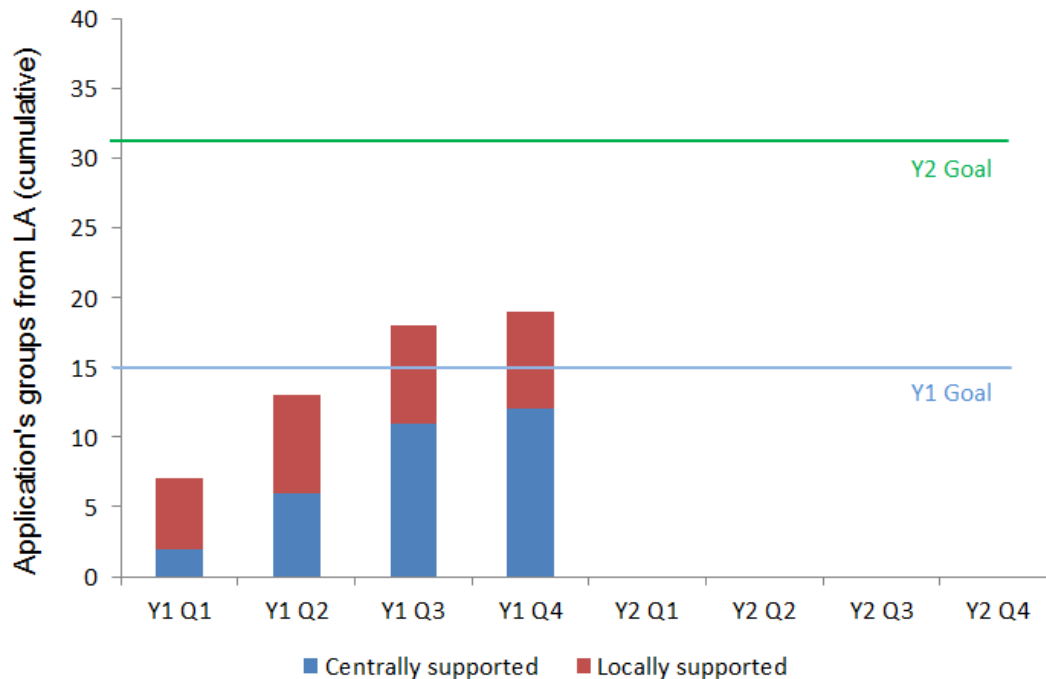


Figure 2: Number of application groups from Latin America

During the Y1Q1, two groups were centrally supported (two high energy physics groups from UFRJ - Auger and LHCb). In Y1Q2, four groups were centrally supported (1 biochemistry - UFRJ, 1 industrial - CEFET, 2 high energy physics - UNIANDES and UNAM). In Y1Q3, five groups were centrally supported (3 biomedicine - UNIANDES, UNIVALLE, UdelaR, 1 physics - UFRJ, 1 astrophysics - UFRJ). Other groups are locally supported (directly by a GISELA partner) by CIEMAT in Spain, UPORTO in Portugal, UNIANDES in Colombia and CEFET-RJ in Brazil. In Y1Q4, one group from UNAM working with the bioinformatics Bowtie¹³ application was centrally supported.

3.2.1. Usage statistics

The use of the computing resources varies from time to time, depending on the set of researchers using GISELA's grid. Data challenges sporadically organised by some VRC, newcomers bringing new experiments and training events like Grid Schools are the "boosters" of the infrastructure.

Figure 3 depicts the CPU-hours consumed by each VO supported by GISELA. One can note that the HEP VRC is the most demanding one, as it the case of the EGI e-Infrastructure with the intense activity of the LHC experiments.

¹³ <http://bowtie-bio.sourceforge.net/index.shtml>

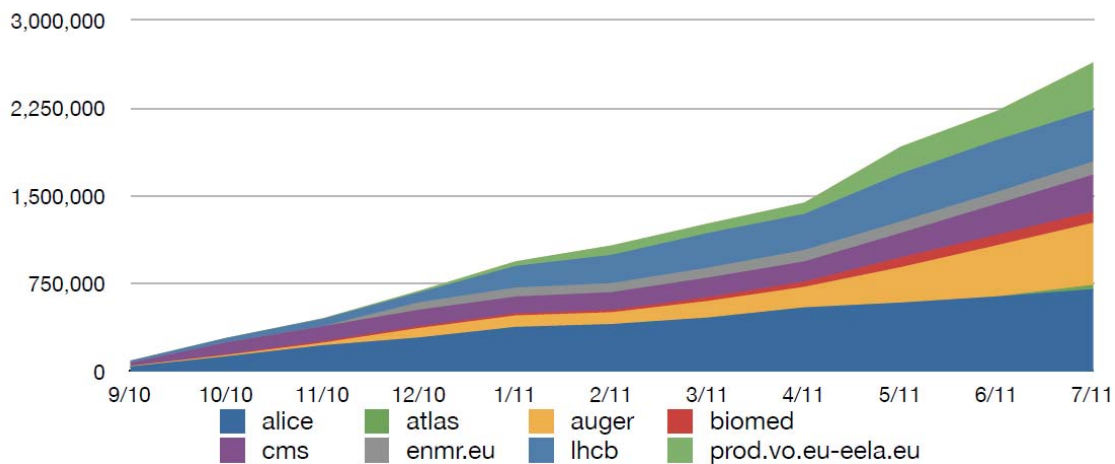


Figure 3: Cumulative CPU-hours (normalized) grouped by VOs

Figure 4 compares the usage made by each VO over the first year of the project.

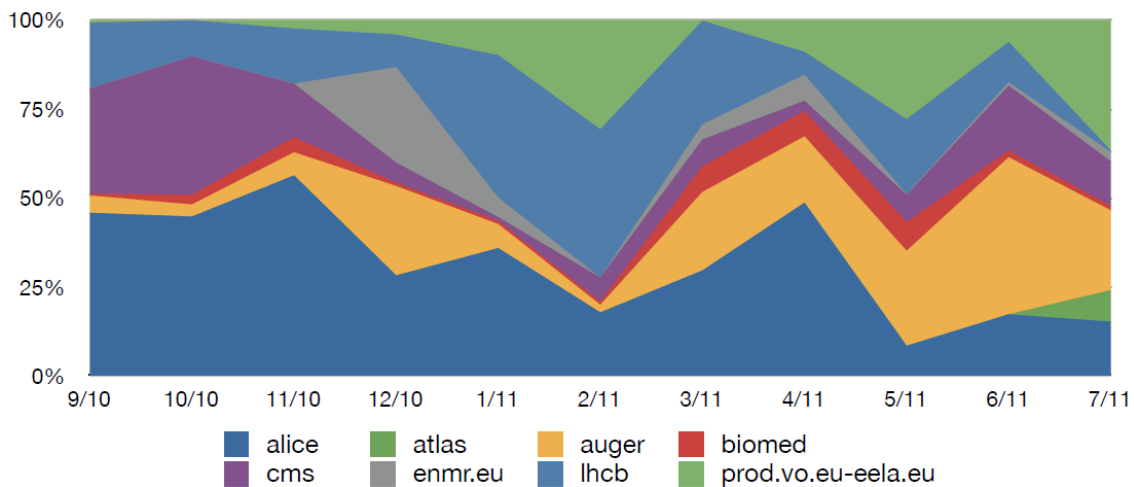


Figure 4: Monthly utilization by each VO (percentage)

3.3. APPLICATIONS PORTFOLIO

The information about the legacy applications inherited from the former EELA and EELA-2 projects can be found in the URL <http://applications.gisela-grid.eu>. Applications identified as belonging to the status “S4” or “S5” – according to the development life-cycle described in http://applications.gisela-grid.eu/app_status.php?l=21 – are those that have being ported to the GISELA grid infrastructure.

It does not mean that all ported applications are currently running on the infrastructure on a production basis. That list should be interpreted as an “available application list”, which can be requested by the GISELA users whenever they need.

All the GISELA applications belonging to the status “S5” have been inserted in the new world-wide applications databases provided by the European Grid Initiative (EGI). The EGI applications database is available at <http://appdb.egi.eu7>.

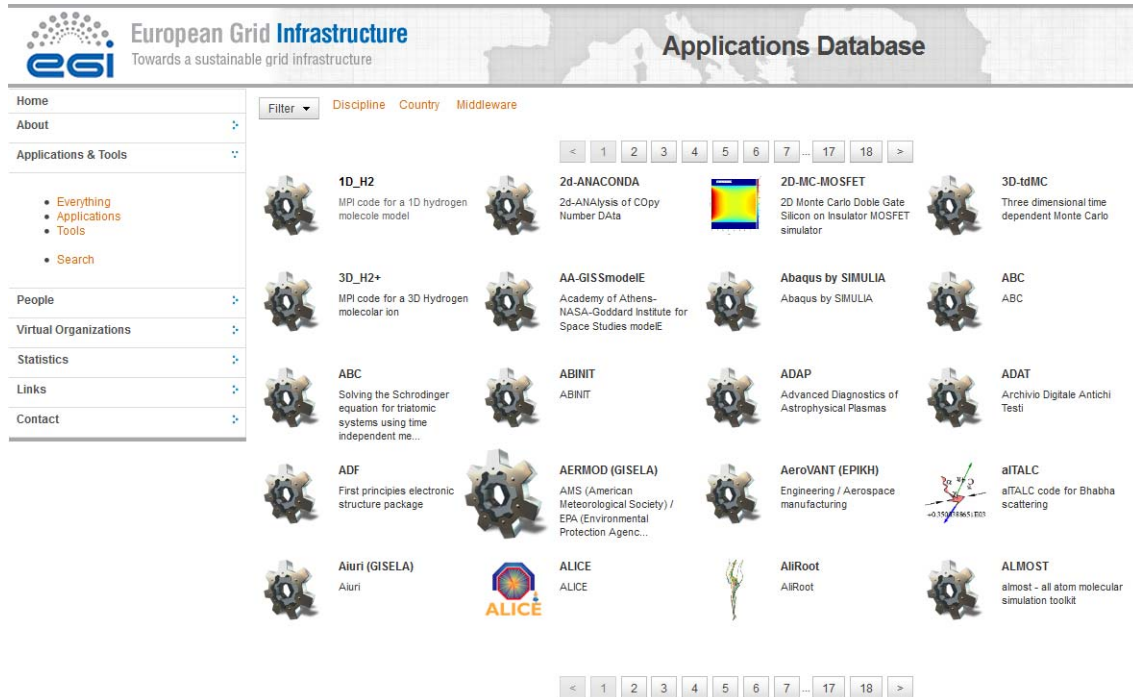


Figure 5: The EGI Applications Database (see <http://appdb.egi.e/>)

Table 4 shows all the GISELA legacy applications in status “S5” providing, for each of them, the link to the related EGI database record.

Table 4: GISELA Applications belonging to status "S5"

Application	Scientific Domain	Country	EGI Applications DB
AERMOD	Earth Sciences	Cuba	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTA
AeroVANT	Engineering	Argentina	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODI
Aiuri	Computer Science and Mathematics	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTE

Application	Scientific Domain	Country	EGI Applications DB
BiG (Blast)	Bioinformatics / Genomics	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTI
BioMD	Life Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODM
bioNMF	Bioinformatics / Genomics	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTM
BRAMS	Earth Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTQ
CAM	Earth Sciences	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD0yODM
CardioGrid Portal	Life Sciences	Argentina	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTU
Cinefilia	Computer Science and Mathematics	Italy / Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTY
CIS - Classification of Satellite Images with neural networks	Earth Sciences	Ecuador	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTc
CROSS-Fire	Civil Protection	Portugal	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01OTk
DicomGrid	Life Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDA
Dist-SOM-PORTRAIT	Bioinformatics / Genomics	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDE
DKEsG	Fusion	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01NjQ
FAFNER2	Fusion	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD0yODQ
fMRI	Life Sciences (e-health)	Portugal	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDI
G-HMMER	Bioinformatics / Genomics	Colombia	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDM
GAMOS	Life Sciences	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDQ
gCSMT	Earth Sciences	France	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD0yNzQ

Application	Scientific Domain	Country	EGI Applications DB
GenecodisGrid	Bioinformatics / Genomics	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDU
Grid Bio Portal	Bioinformatics / Genomics	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDY
GRIP - Grid Image Processing for Biomedical Diagnosis	Life Sciences	Chile	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODU
GROMACS	Life Sciences (Chemistry)	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDc
gRREEMM	Engineering	Cuba	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MDg
HeMoLab	Life Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODY
Industry@Grid	Engineering	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01NjU
InvCell	Life Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MTE
jModelTest	Life Sciences	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MTI
PILP	Computer Science and Mathematics	Portugal	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MTQ
ProtozoaDB	Life Sciences	Brazil	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODc
Seismic Sensor	Earth Sciences	Mexico	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD01ODk
WAM	Earth Sciences	Ireland	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD02MTU
WRF	Earth Sciences	Spain	http://appdb.egi.eu/?p=L2FwcHMvZGV0YWlscz9pZD0yODI

Table 5 lists GISELA legacy applications in status “S4” providing, for each of them, the link to the related GISELA applications database record. When these applications reach the status “S5”, they will be inserted on the EGI applications database too.

Table 5: GISELA Applications belonging to status "S4"

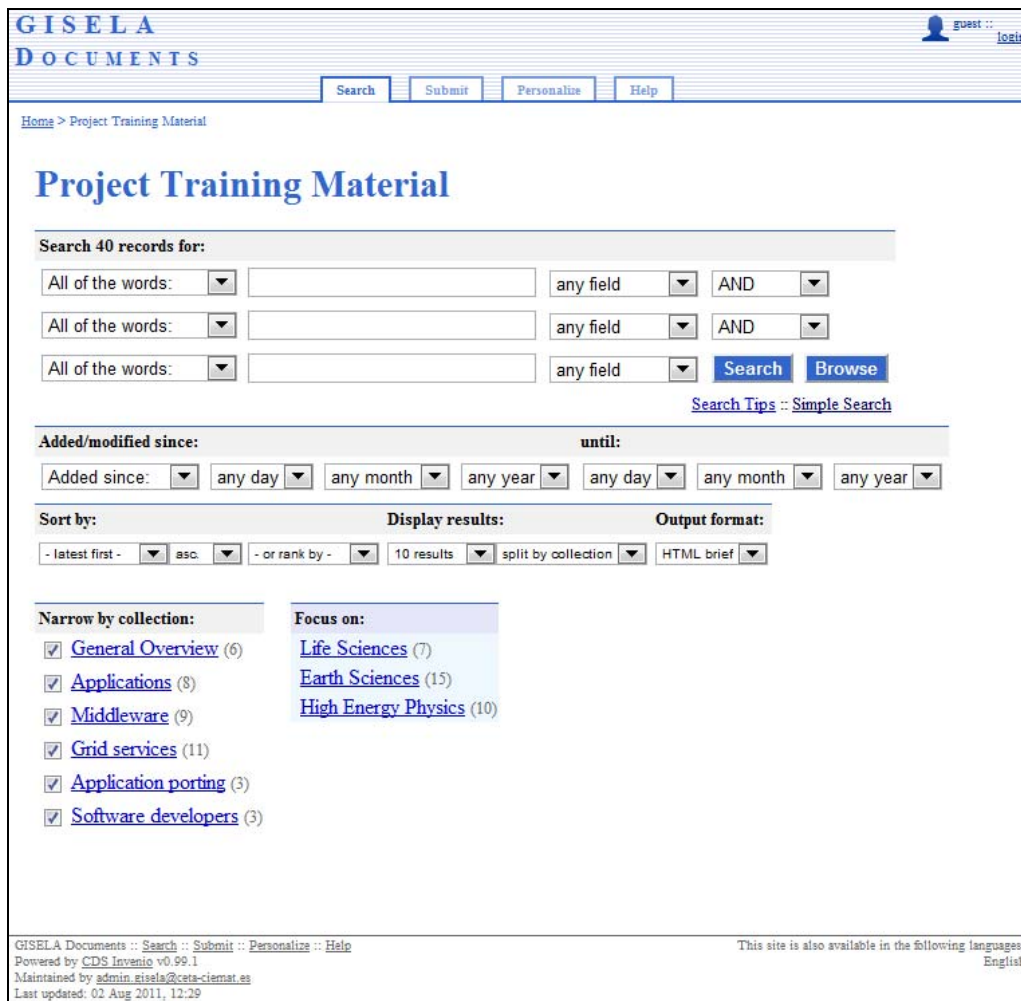
Application	Scientific Domain	Country	GISELA Applications DB
APPF	Computer Science and Mathematics	Mexico	http://applications.gisela-grid.eu/application_details.php?l=20&ID=81
C/CATT-BRAMS	Earth Sciences	Chile / Brazil	http://applications.gisela-grid.eu/application_details.php?l=20&ID=51
CTSAE	Life Sciences	Mexico	http://applications.gisela-grid.eu/application_details.php?l=20&ID=83
DistBlast	Bioinformatics / Genomics	Brazil	http://applications.gisela-grid.eu/application_details.php?l=20&ID=49
DRI/Mammogrid	Life Sciences (e-health)	Spain	http://applications.gisela-grid.eu/application_details.php?l=20&ID=32
eIMRT	Life Sciences (e-health)	Spain	http://applications.gisela-grid.eu/application_details.php?l=20&ID=21
GrEMBOSS	Bioinformatics / Genomics	Mexico	http://applications.gisela-grid.eu/application_details.php?l=20&ID=11
Heart Simulator	Life Sciences	Brazil	http://applications.gisela-grid.eu/application_details.php?l=20&ID=12
InvTissue	Life Sciences	Brazil	http://applications.gisela-grid.eu/application_details.php?l=20&ID=62
MAVs-Study	Engineering	Argentina	http://applications.gisela-grid.eu/application_details.php?l=20&ID=72
META-Dock	Bioinformatics / Genomics	Mexico	http://applications.gisela-grid.eu/application_details.php?l=20&ID=55
Portal de Porticos	Engineering	Venezuela	http://applications.gisela-grid.eu/application_details.php?l=20&ID=44
SATCA	Earth Sciences	Mexico	http://applications.gisela-grid.eu/application_details.php?l=20&ID=69
SEMUM3D	Earth Sciences	France	http://applications.gisela-grid.eu/application_details.php?l=20&ID=18

In addition to the applications listed in the tables above, applications from other projects can be accepted to run on GISELA infrastructure. For example, according to the MoU between GISELA and EGI-InSPIRE, VRCs within EGI could exploit GISELA resources to run their applications.

3.4. USER SUPPORT

3.4.1. Documentation

Every training material produced during first year was reviewed, tagged and made available through the GISELA Documents Server¹⁴, where users, administrators, developers and newcomers are able to search for useful documentation in a centralized and effective way.



The screenshot shows the GISELA Documents search portal. The page title is "Project Training Material". The search interface includes a search bar with "Search 40 records for:" and three input fields for search terms. Below the search bar are filters for "Added/modified since" and "until" with dropdown menus for time intervals. There are also options for "Sort by" (latest first, asc, - or rank by -), "Display results" (10 results, split by collection), and "Output format" (HTML brief). On the left, there is a "Narrow by collection" section with checkboxes for "General Overview (6)", "Applications (8)", "Middleware (9)", "Grid services (11)", "Application porting (3)", and "Software developers (3)". On the right, there is a "Focus on:" section with links for "Life Sciences (7)", "Earth Sciences (15)", and "High Energy Physics (10)". The footer contains the text: "GISELA Documents :: Search :: Submit :: Personalize :: Help", "Powered by CDS Invenio v0.99.1", "Maintained by admin.gisela@ceta-ciemat.es", "Last updated: 02 Aug 2011, 12:29", and "This site is also available in the following languages: English".

Figure 6: Project Training Material search portal

¹⁴ <http://documents.gisela-grid.eu/>

As shown in Figure 7, 40 training documents (user guides, installation guidelines, FAQs, etc) have been uploaded to the GISELA Documents Server over the first year. Most of these documents were produced by tutors of, and attendants to, local workshops and other training events co-organised by GISELA, CHAIN and EPIKH (see “Tutorials and Schools” section on GISELA Events Server¹⁵). However, a few others (7) were co-produced with other projects before the beginning of GISELA.

Training material has been organised in 6 “narrow by” collections in order to ease selective searches and categorize documents according to the interests of multiple user profiles:

- **General overview** – documents for newcomers and non-technical individuals.
- **Applications** – documents describing scientific applications and algorithms that may be of special interest to target VRC members.
- **Middleware** – documentation related to any middleware that composes the GISELA grid infrastructure.
- **Grid services** – documents and presentations describing RC and ROC services from grid site administrator point of view.
- **Application porting** – contains a FAQ document introducing the process of adapting scientific applications to grid; it also contains presentations describing several application porting processes and results produced during joint GISELA/CHAIN/EPIKH Schools for Application Porting.
- **Software developers** – training documents relevant to developers wanting to integrate grid resources within portals, tools and applications.

Apart from these some kind of static categories, “focus on” virtual collections have been provided for each targeted VRC: High Energy Physics (HEP), Earth Sciences and Life Sciences. When accessing these virtual collections, users can get documents of any kind (not only training material, but trip reports, meeting minutes, publications and so on), produced by GISELA and relevant to the focused VRC.

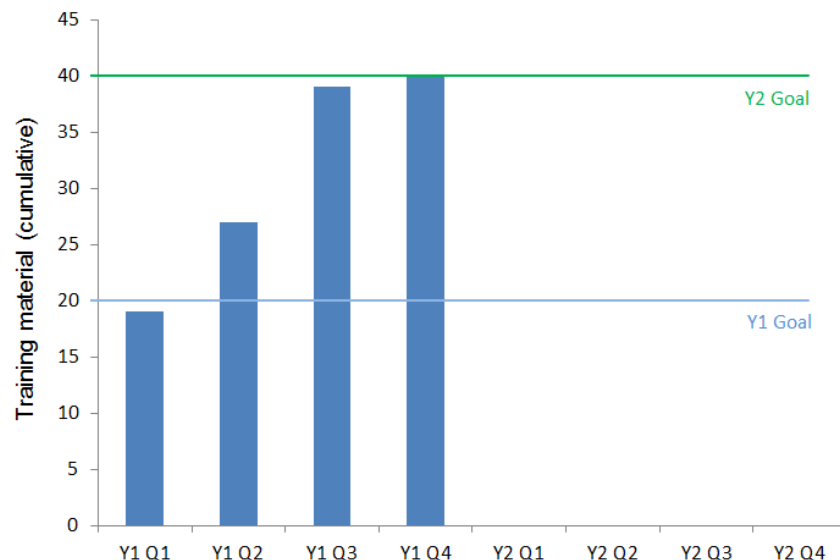


Figure 7: Number of self-training material made available through Documents Server

¹⁵ <http://indico.gisela-grid.eu>

3.4.2. Publications

Up to 30 GISELA related publications can be consulted at the “Publications” area in GISELA Web Page¹⁶. From these, 5 papers related to the activity of WP3 were published during this year, so the quality metric committed in the DoW was successfully accomplished.

- BARBERA, Roberto; BRASILEIRO, Francisco; BRUNO, Riccardo; CIUFFO, Leandro and SCARDACI, Diego. **Supporting e-Science Applications on e-Infrastructure: Some Use Cases from Latin America**. Book Chapter: “*Grid Computing. Towards a Global Interconnected Infrastructure*”, Editor "Springer", pages 33-55, July 2011;
- ISEA, Raul; MONTES, Esther; RUBIO-MONTERO, Antonio; MAYO, Rafael. **State-of-art with PhyloGrid: Grid computing Phylogenetic studies on the EELA-2 Project infrastructure**, *Computer Communications and Networks 2011*, SPRINGERLINK, 2011 - <http://www.springerlink.com/content/n100450864364113/>
[Abstract in the GISELA Document Server](#);
- PINA, Antonio; ESTEVES, Antonio; PUGA, Joel; and OLIVEIRA, Vitor. **A Geographical Information System for wild fire management**. 5th Iberian Grid Infrastructure Conference (Ibergrid'2011), 8th -10th June 2011, Santander (Spain) - <http://documents.gisela-grid.eu/record/211?ln=en>;
- GARCÍA, Sebastián; ITURRIAGA, Santiago; NESMACHNOW, Sergio (Universidad de la República, Uruguay). **Scientific computing in the Latin America-Europe GISELA Grid infrastructure**. EProceedings of the High-Performance Computing Latin America Symposium (HPCLatAm2011), Cordoba, Argentina, 2011.
Draft version: <http://www.fing.edu.uy/inco/grupos/cecal/hpc/publications/sciomp-GISELA.pdf>;
- ANDRONICO, Giuseppe; ARDIZZONE, Valeria; BARBERA, Roberto; BECKER, Bruce; BRUNO, Riccardo; CALANDUCCI, Antonio; CARVALHO, Diego; CIUFFO, Leandro; FARGETTA, Marco; GIORGIO, Emidio; et al. **e-Infrastructures for e-Science: A Global View**, *Journal of Grid Computing*, SPRINGERLINK Online First™, 24th March 2011 - <http://www.springerlink.com/content/81268575123q1p49/>
[Abstract in the GISELA Document Server](#)

3.4.3. Training events

Two training events were jointly organised with the EPIKH and CHAIN projects at first quarter of the year. Additionally, two workshops were organised by UNAM at Hermosillo and Guaymas (Sonora Mexico). Table 6 lists the training events organised so far.

Table 6: List of training events organised in the first year

Event	Date	Venue
1 st Joint GISELA/EPIKH School for Application Porting and Grid Site Administrators	November 2010	Mexico City, Mexico

¹⁶ http://www.gisela-grid.eu/index.php?option=com_content&view=article&id=43&Itemid=41

Event	Date	Venue
2 nd Joint CHAIN/GISELA/EPIKH School for Application Porting and Grid Site Administrators	November 2010	Valparaiso, Chile
Workshop: <i>Uso de Infraestructuras Grid Computacional en Aplicaciones de Ciencia y Tecnología</i>	8 th - 10 th June 2011	Hermosillo, Mexico
Workshop: <i>Uso de Infraestructuras Grid Computacional en Aplicaciones de Ciencia y Tecnología</i>	13 th - 17 th , June 2011	Guaymas, Sonora, Mexico

These four training schools and workshops covered the following disciplines:

- Accessing to grid computing and storage facilities;
- Administration of Resource Centres (RCs);
- Dissemination of already ported applications relevant to HEP, Life Science and Earth Science VRCs;
- Porting of some new applications required by members of Life Science VRC.

Figure 8 shows the accomplishment of the quality metric related to the number of training events. One or two more joint training events in Latin America are expected to be organised until the end of the project.

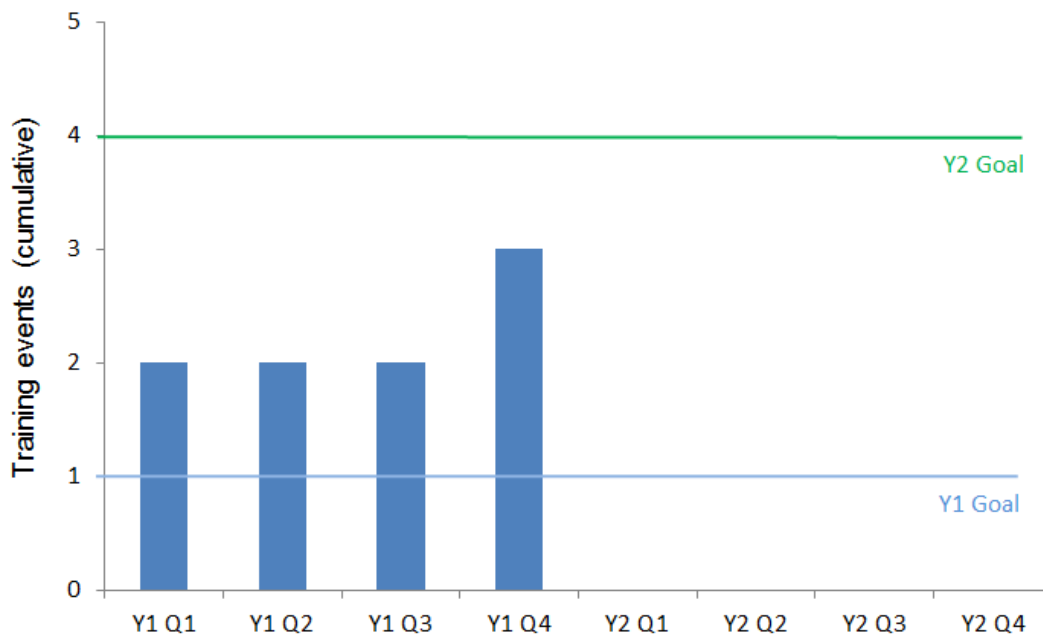


Figure 8: Number of Training Events

4. SCIENCE GATEWAY FOR INDUSTRIAL APPLICATIONS

During its first year, GISELA decided to prospect the use of Science Gateways to potentially increase the e-infrastructure usage. Science Gateways are considered as valid and innovative tools to increase Grid adoption and usage. Hiding the complexity of the Grid environment could indeed allow an easy access to e-Infrastructures to large Virtual Research Communities reducing the skills needed today to fully exploit them.

In collaboration with WP6, we decided to design and develop a Science Gateway for the Industrial Applications. We chose a pilot application to be integrated in this Science Gateway, Industry@Grid (see http://applications.gisela-grid.eu/application_details.php?l=20&ID=14) and we started to collaborate with its application developers team. This Science Gateway will be developed using international standard as JSR 168¹⁷ e JSR 286¹⁸ for Web Portal development and the Simple API for Grid Applications (SAGA) Core API, a high level, application-oriented, software library for Grid application development specified by the Open Grid Forum. SAGA allows to create a unique interface towards different middleware stacks and makes Scientific Gateways able to exploit resources coming from different Grid worlds.

Figure 9 depicts the reference model we adopted to develop the Science Gateway for the Industrial Applications. The model is generic and could be used for other applications domains or VRCS. Then, in case of a good feedback from the users communities, we could decide to develop new Science Gateways for some other application domain or VRC.

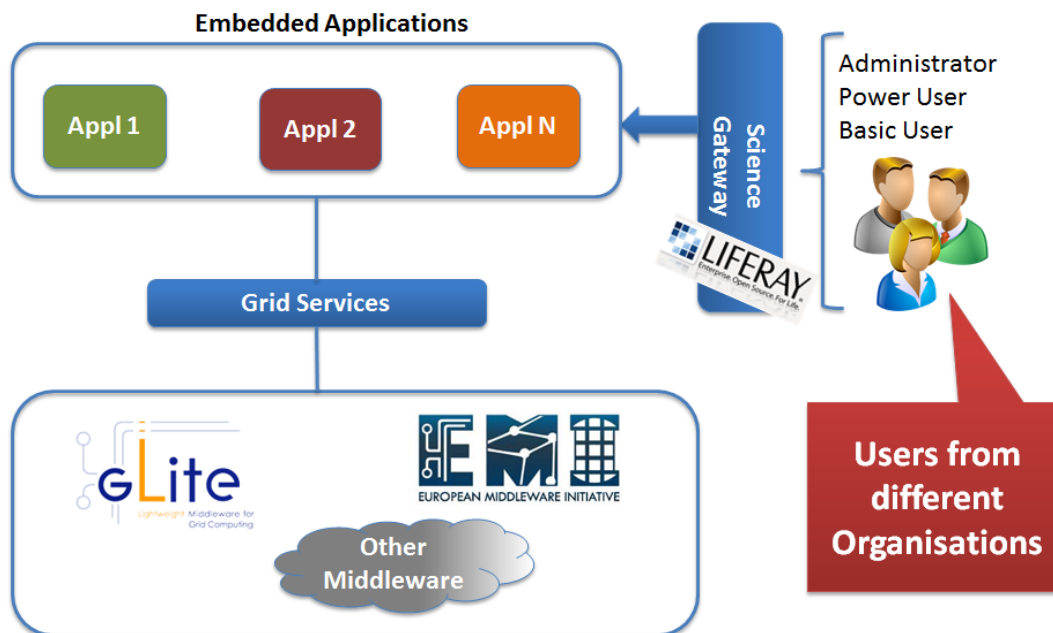


Figure 9: Reference model of the Science Gateway for the Industrial Applications

¹⁷ <http://www.jcp.org/ja/jsr/detail?id=168>

¹⁸ <http://www.jcp.org/ja/jsr/detail?id=286>

5. HUMAN EFFORT

The current human resources allocated to WP3 are listed in Table 7 and Table 8. The data have been extracted from the GISELA timesheets system.

Table 7: Human effort allocated to WP3 management

Name	Role	Institution
Diego Carvalho	WP3 Manager (until M11)	CEFET/RJ
Leandro N. Ciuffo	WP3 Manager pro tem (M11 and M12)	RNP
Rafael Mayo	Deputy Manager	CIEMAT
Diego Scardaci	Task Leader 3.2	INFN - Catania
Guillermo Diaz Herrero	Task Leader 3.3	CIEMAT
Dago Bedoya	Clara TT member allocated to WP3	RENATA
Diego Alberto Rincón Yáñez	Clara TT member allocated to WP3	RENATA

Table 8: List of application's developers/supporters

Name	Institution
Abelardo Roddriguez	UAM
Alexander Herrera	UNIANDES
Alina Roig	CUBAENERGIA
Anaís Hernández Ros	CUBAENERGIA
Angelines Alberto	CIEMAT
Angelines Alberto	CIEMAT
César Fernandez	UTFSM
David Méndez	UNIANDES

Name	Institution
Enrique González	UNIANDES
Fredy Juarez	UAEM (UNAM)
Gabriel Usera	UdelaR
Genevieve Moguilny	CNRS
Gonzalo Rodriguez	UdelaR
Guilherme Baptista	UFCG
Harold Castro	UNIANDES
Henry Ricardo Mora	CUBAENERGIA
Jesus Cruz	UNAM
Jhoanna Serpa	CEDIA
Jorge Blanco	CIEMAT
José Luis Garza Rivera	UNAM
Julio López	UFRO
Luiz Rossi de Souza	UFRJ
Maikel Isover Domínguez García	CUBAENERGIA
Manuel Rodriguez	CIEMAT
Mario Villamizar	UNIANDES
Mauro Canabe	UdelaR
Moises Hernandez Duarte	UNAM
Natalia Garces	UNIANDES
Paola Arce Azócar	UTFSM
Paulina López	REUNA
Querube Urriola	CIDETYS
Rafael Barbastefano	UFRJ
Raquel Pezoa	UTFSM

Name	Institution
Santiago Iturriaga	UdelaR
Sebastian García	UdelaR
Sergio Nesmachnow	UdelaR
Tarciso de Oliveira Filho	UFCG
Villie Morocho	CEDIA

6. OPEN ISSUES AND / OR DEVIATIONS FROM THE WORK PLAN

On M11, Diego Carvalho assumed a new position in his institution, leading him to step down from his role of WP3 manager. Leandro Ciuffo (RNP) was appointed to become the WP3 Manager pro tempore.

The slow start of the CLARA Transition Team person in WP3, which has been nominated late in the course of the project (Q3), can be mentioned as a deviation. Unfortunately, even after such a nomination, the CLARA TT is not yet fully engaged within the WP3 activities.

Also, the coordination of VRC recruitment from CLARA communities could not be well coordinated between the WP3 management and CLARA & NRENs due to the different initial vision of potential new communities.

With regards to the use of the infrastructure, one may interpret that it is still low. Nevertheless, we should bear in mind that some Applications usually exploiting the Grid have their own characteristics. Thus, it is important to point out that the major Grid users, i.e. LHC experiments and Biomed VO are already supported by GISELA, i.e. the project has guaranteed in principle that a big submission of jobs related to these disciplines will be done successful. On the other side, some LA sites cannot comply with some specific requirements that, for example, LHC production has, so their jobs cannot be processed in those sites and a limitation in the use of the infrastructures rises. This observation could be extended to other applications (DKESG, for example) which need 64 bits calculations due to the required precision in some part of their calculations. Also, some Earth Science models rely on MPI for execution, what can reduce the number of available grid sites if this support is not widespread.

Hence, at the same time that the heterogeneity of resources (and middleware flavours) can be a value in a continental grid infrastructure such as GISELA, it also can plays against the availability of resources for each VRC or application. The complete saturation of the infrastructure on a regular basis is a target rather difficult to be achieved. The WP3 work shall focus on disseminating GISELA's grid as a tool for e-science, prospecting new VRCs and be ready to support research groups whenever they need help.

7. PLANS FOR THE NEXT REPORTING PERIOD

A new WP3 “VRC-driven” website will be created, replacing the out-dated content available at <http://applications.gisela-grid.eu/>. The new application web site will use the last web technologies as Liferay (see <http://www.liferay.com/>) and international standard for Web Portal development as JSR 168 and JSR 286.

At least two more training events are planned to be organised during the second year in cooperation with EPIKH, CHAIN, and DEGISCO¹⁹ possibly. Tutors from member institutions attending these training events will be in charge of generating more self-training material, or improving those created in previous events. Every training material produced will be reviewed and made available through the GISELA Documents Server.

The possibility of hosting a new industry-related application will be pursued.

As aforementioned, the Brazilian National Institute for the Research of the Amazon-Region (INPA) has been approached by WP3. We plan to support INPA to evaluate the WRF4G model ability to represent the inter-annual variability in precipitation and circulation patterns of wet season in central Amazon from 1988 to 1999. The estimated time for running such a simulation is 15 days. Also, it is expected to consume 1TB of storage

In previous section, some issues related to the use of the infrastructure have been mentioned. WP3 has not only detected them, but some actions have been taken in order to minimize their negative impact. Thus, several researchers have planned jointly with WP3 management a set of applications that would increase the use of the e-infrastructure. Specifically, several biological cases related to models of substitution will be studied by means of jModelTest and a deep research about the Neutral Beam Injection facility of TJ-II stellerator will be also carried out by FAFNER2 (both applications already count on the GISELA portfolio).

In addition, the Bioinformatics Group from the Universidad Nacional de Colombia is expected to be requesting certificates to the newly created CA in Colombia (UNIANDES) in a very short time. Since they will be affiliated to appropriate GISELA e-Infrastructure VO(s), a higher exploitation of the e-Infrastructure resources committed to Life Science VRC will be produced.

Concerning the engagement from CLARA TT, WP3 is working with WP1 to mitigate their lack of participation, as it is essential for planning the handover of the user support to RedCLARA at the end of the project.

7.1. DEVELOPMENTS AT UNIANDES (COLOMBIA)

The local support team at UNIANDES has been very proactive on deploying bioinformatics applications jointly with the university’s biology department. Four bioinformatics application suites have been installed so far on a local implementation: NCBI BLAST version 2.2.20, HMMER version 2.3.2, InterProScan version 4.6 and mpiBLAST version 1.6.0.

Since mpiBLAST requires the use MPI implementation, the MPICH-2 version 1.2.7 was installed. Twenty-one LONI²⁰ Modules were created for the application suites: 6 for NCBI BLAST (BLASTn, BLASTp, BLASTx, tBLASTn, tBLASTx and MEGABLAST); 3 for HMMER (Build, Search and Calibrate); 11 for InterProScan (BlastProdom, Coils, Gene3D, PIR, Panther, Pfam, SEG, SMART, SuperFamily, TIGRfam and fPrintScan) and 1 for mpiBLAST.

¹⁹ <http://degisco.eu/>

²⁰ A workflow tool originally designed for neurosciences.



1ST YEAR ACTIVITY HIGHLIGHTS ON VRCS AND USER SUPPORT

Document Full Name

GISELA-D3.2-v1.8

Date: 26/08/2011

Local researchers are now able to easily run these bioinformatics applications using graphical user interfaces and the above workflows, which are executed on a local opportunistic distributed infrastructure.

For the next reporting period, UNIANDES is planning to integrate all this work on the GISELA e-infrastructure, making available their bioinformatics workflows to the biology community within the project.

8. CONCLUSIONS

WP3 is fulfilling what has been committed in the DoW by offering appropriate support at VRC level. As a result the success metrics set for year-1 on the numbers of supported VRCS, supported groups in Latin America, Training events and material as well as on the number of papers published, have all been met. The WP4 work on enabling the GISELA resources to accept computing jobs from VOs such as “WeNMR”, “biomed” or “auger” is of fundamental importance for this accomplishment. Also, thanks to the continuation of the legacy “prod.vo.eu-eela.eu” VO, more than 50 individual applications can still run on the GISELA infrastructure.

However, as aforementioned, it does not mean that all previously ported applications are currently running on the infrastructure on a production basis. Most of those applications are no longer used by the research group who developed it. Those applications now compose a “directory of tools” that can be requested by users interested to learn from them.

The work plan for year-2 is clearly defined. It focuses on completing the VRC support environment (e.g. with the creation of specific Website), organising two more training events in collaboration with other projects as found very convenient and on prospecting new User communities from various sectors as climate change studies, bioinformatics but also from RedCLARA communities coming from the social and cultural horizon.