



Design Study



European Grid Initiative Design Study (EGI_DS)

THE EGI KNOWLEDGE BASE – FIRST SNAPSHOT

WORK PACKAGE 2
EGI REQUIREMENTS CONSOLIDATION

EU DELIVERABLE: D2.2

Document identifier:	EGI_DS-WP2-D2.2-250108-v0.1.doc
Date:	19/05/2008
Work Package:	WP2: EGI Requirements Consolidation
Lead Partner:	INFN
Document status:	FINAL
Document link:	http://web.eu-egi.org/owl/browse.php?sess=0&parent=68&expand=1&fileid=164

Abstract: This document describes the website known as the *EGI Knowledge Base*, as of month 5 of the EGI_DS project. The site itself – at <http://knowledge.eu-egi.org> – is the WP2 prototype deliverable D2.2. As with any website, the EGI Knowledge Base is a dynamic entity whose content evolves with the needs of the community it serves.



Copyright notice:

Copyright © Members of the EGEE-II Collaboration, 2006.

See www.eu-egee.org for details on the copyright holders.

EGEE-II (“Enabling Grids for E-sciencE-II”) is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 6th Framework Programme. EGEE-II began in April 2006 and will run for 2 years.

For more information on EGEE-II, its partners and contributors please see www.eu-egee.org

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 EGEE-II Collaboration 2006. See www.eu-egee.org 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 are published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS “AS IS” AND ANY EXPRESS 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 EGEE-II 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.

Trademarks: EGEE and gLite are registered trademarks held by CERN on behalf of the EGEE collaboration. All rights reserved”

Delivery Slip

	Name	Partner/Activity	Date	Signature
From	Diana Cresti	INFN/WP2	26/01/08	
Reviewed by	Leif Laaksonen Michael Wilson	CSC STFC		
Approved by	PEB			

Document Log

Issue	Date	Comment	Author/Partner
0.1	26/01/08	First draft and Table of Contents draft	Diana Cresti / INFN
0.2	10/04/08	First review	Leif Laaksonen
0.3	28/04/08	Second draft, incorporating response to review comments	Diana Cresti
0.4	28/04/08	General revisions (e.g. Terminology text)	Diana Cresti

Document Change Record

Issue	Item	Reason for Change



TABLE OF CONTENTS

1. INTRODUCTION.....	4
1.1. PURPOSE.....	4
1.2. DOCUMENT ORGANISATION.....	4
1.3. TARGET AUDIENCE.....	4
1.4. REFERENCES.....	4
1.5. DOCUMENT AMENDMENT PROCEDURE.....	4
1.6. TERMINOLOGY.....	5
2. EXECUTIVE SUMMARY.....	6
3. DESCRIPTION OF THE INSTALLATION.....	6
4. GENERAL STRUCTURE AND CONTENTS.....	7
4.1. MAIN MENU.....	7
4.1.1. EGI Functions.....	8
4.1.2. Use Cases.....	8
4.1.3. European Projects.....	8
4.1.4. Application Areas.....	9
4.1.5. Legal Options.....	9
4.1.6. The Internal area.....	9
4.2. THE NGI DESCRIPTIONS AREA.....	9
4.3. INTEGRATION WITH THE E-IRG KNOWLEDGE BASE.....	11
5. ANNEXES.....	12

TABLE OF TABLES

Table 1: Table of references.....	4
------------------------------------------	----------



1. INTRODUCTION

1.1. PURPOSE

This document describes the website known as the *EGI Knowledge Base*, as of month 5 of the EGI_DS project. The site itself – at <http://knowledge.eu-egi.org> – is the WP2 deliverable D2.2; hence this is a software deliverable, not a document deliverable.

What follows is thus a brief description of the software tool(s) used in creating this site, its structure and contents as of the date of writing of this document, and some discussion of the plans and rationales motivating the choices being made in relation to this site. As with any website, the EGI Knowledge Base is a dynamic entity whose content evolves with the needs of the community it serves.

The purpose of the EGI Knowledge Base (EGI KB) is

- to document the status and evolution of the European National Grid Initiatives (NGIs),
- to collect information about current e-Infrastructure activities and projects in Europe,
- to share the ongoing work done by the EGI Design Study team, and
- to provide a real time feedback loop between the team and the NGI representatives.

Thus the EGI KB was initially conceived not just as a source of general information on the European e-Infrastructure, but also as a working tool and a communication channel for the EGI_DS team and the NGIs – a tool that could in principle replace some of the communications that are generally conducted via email and provide a structured space for important documentation. From the description of the site's content in section 4 (and by checking the site itself), one can see that the EGI KB is being used in part for the purposes it was intended, but there is certainly room for improvement. In section 4.2 and 4.3 we briefly discuss steps being taken to improve the use and usefulness of the site.

1.2. DOCUMENT ORGANISATION

This document provides a technical description of the EGI Knowledge Base in section 3, and a description of its structure and contents in section 4.

1.3. TARGET AUDIENCE

This document is a public deliverable intended to inform all the EGI related stakeholders, as well as readers not directly related with EGI (e.g. from other geographical areas) and the general public.

1.4. REFERENCES

[This subsection provides a complete list of all documents referenced elsewhere in the document.]

Table 1: Table of references

R 1	The EGI Knowledge Base: http://knowledge.eu-egi.org
R 2	

1.5. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the author, Diana Cresti, at diana.cresti@pd.infn.it.



1.6. TERMINOLOGY

This subsection provides the definitions of terms, acronyms, and abbreviations required to properly interpret this document. A basic assumption that has been also agreed recently is that the term EGI (European Grid Initiative) refers to both the EGI organisation and NGIs, while the EGI organisation refers to the EGI headquarters, where the team with central responsibility is located, including the management structure. A useful glossary of terms can be also found under <http://www.eu-egi.org/internal/glossary>.

Glossary

2. EXECUTIVE SUMMARY

The EGI / NGI environment is as uniquely complex as Europe itself. This poses a challenge for those who wish to give adequate representation to the heterogeneous realities within this environment and to their convergent aspects. The best speakers for these realities are the NGI representatives themselves, as well as the e-Infrastructure specialists identified by the EGI_DS team and the NGI representatives.

The desideratum of providing a single location where the EGI / NGI environment can be described, while at the same time giving a voice to all the main stakeholders in an efficient and non-centralised manner, leads naturally to the concept of a **multi-author environment**, which in turn leads to the choice of a stable and user-friendly content management system, of which a wiki-type tool is probably the best example.

The most powerful such tool is MediaWiki – <http://www.mediawiki.org>. This wiki not only provides discussion pages for every “article” (web page) in a site, its library of extensions is unparalleled, ranging from graphic enhancements such as image map support, to sophisticated tools for extracting survey and statistical data from anywhere on the site. Furthermore, it includes a system of permissions which allows the establishment of private as well as public areas in the site. Finally, this tool has a very solid and intuitive “history” function which allows the user to recover earlier versions of a page with just a few clicks.

A potential difficulty with a multi-author environment is that the updating of content cannot be ensured as part of somebody’s “job description”. This is a particular challenge when the “authors” are NGI representatives who have little time for content management. Hence the EGI_DS team has been adding content managers (appointed by the NGI representatives) to the list of authorised users of the EGI KB. The planned integration with the e-IRG knowledge base and other similar resources is also intended to harmonise the process of acquiring information from the NGIs.

3. DESCRIPTION OF THE INSTALLATION

The EGI KB was first installed in August 2007. The members of the EGI_DS team and all the NGI representatives were given accounts, which allowed them to view and edit the contents of the site. On December 16, the KB was opened (for viewing) to the public; editing, however, will remain restricted to registered users. Furthermore, there is an Internal area (see subsection 4.1.6) which is only viewable by registered users.

The MediaWiki package is installed on a GUP server. It requires “recent” versions of PHP and MySQL. As of February 2008, the relevant versions are:

- **MediaWiki:** 1.10.0
- **PHP:** 5.2.5-pl1-gentoo (apache2handler)
- **MySQL:** 5.0.44-log

Several extensions are currently installed, which augment the functionality and user-friendliness of the wiki, in particular:

- *ImageMap* (<http://www.mediawiki.org/wiki/Extension:ImageMap>) which allows client-side clickable image maps using an <imagemap> tag. This extension incorporates image maps into the wiki environment. It can be seen at work in the Knowledge Base main page, where a clickable map of Europe leads to the individual NGI articles by clicking on the respective countries (see also Figure 1).

- *DynamicPageList2* (v 1.6.4) (<http://www.mediawiki.org/wiki/Extension:DynamicPageList>), based on *DynamicPageList*, is a powerful extension that allows the creation of surveys taken from anywhere in the site. A first example of the use of this can be seen at http://knowledge.eu-egi.org/index.php/NGIs_in_numbers, a list that gives numbers of users and resources for each NGI. Note that these lists are like “live” database calls, so when content is updated in one of the source pages, the survey page is updated as well.

For further details see <http://knowledge.eu-egi.org/index.php/Special:Version>.

Another aspect of MediaWiki that enables high-level customisations, including a means to interact with the survey tool, is the creation of templates. For a list of current templates, see <http://knowledge.eu-egi.org/index.php?title=Special%3APrefixindex&from=&namespace=10>.

Finally, one useful feature of the MediaWiki tool is that each “article” of the wiki has a corresponding “discussion” page, accessible by clicking the *discussion* tab at the top of the article. This area allows users to add comments and other feedback about the article. One can see by looking at Wikipedia how useful this area is in the creation of the site’s content. A good example of this is the article on Grid Computing (http://en.wikipedia.org/wiki/Grid_computing), which has a very lively Discussion page (http://en.wikipedia.org/wiki/Talk:Grid_computing).

4. GENERAL STRUCTURE AND CONTENTS

Like other wikis, the MediaWiki tool allows great freedom in the creation of new content. A new page can be created simply by adding a “link” to some preexisting content. Thus a site built with this tool is essentially an unstructured, relational entity. There are, however, a number of ways in which content can be found:

- The wiki “search engine” (see search box to the left under main navigation menu) is very efficient, and will find main articles with the title matching the string searched (e.g. countries, projects), or if such articles don’t exist will return the results of a text search on the string.
- By clicking on *Special pages* in the toolbox menu (under the search box), one can find a set of links that generate lists of pages in the wiki that fit certain criteria – e.g. New pages, Long pages, Most linked to pages, etc.

The unstructured nature of the MediaWiki tool entails that the users and administrators of the site have the freedom – but also the responsibility – to establish access to important contents in a thematic manner from any relevant “article”. For the EGI KB a general structure was given by modifying the top left sidebar – the *navigation* box – to provide a main menu.

4.1. MAIN MENU

Currently, the main menu contains:

- Main page
- EGI Functions
- Use Cases
- European Projects
- Application Areas
- Legal Options
- Recent Changes

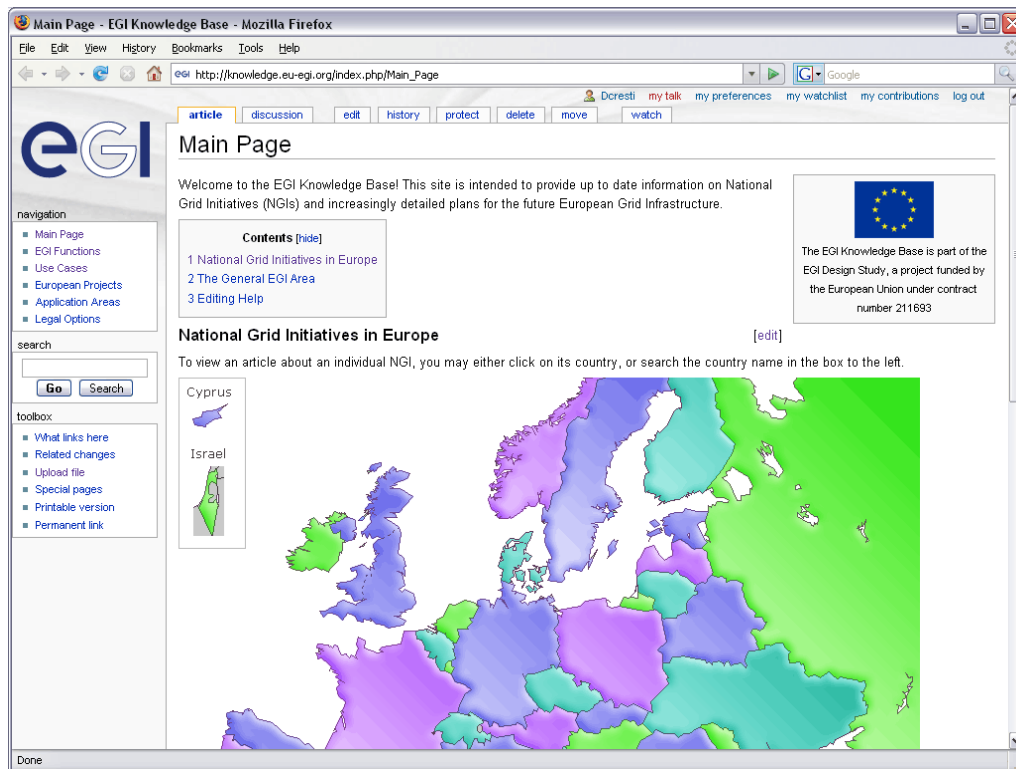


Figure 1: The EGI KB Main Page

In addition, for logged in registered users, there is a further menu item:

- Internal

This area is reserved for the NGI representatives and the EGI_DS team, and is not visible to the public.

In addition to this general menu, the image map on the main page allows easy access to the individual NGIs' articles. The latter can also be accessed from any page in the wiki by typing the country name in the search box to the left. The wiki administrators have created several redirect pages (something that can be done by any logged in user), so that one can for instance search 'uk' and be taken to the United Kingdom NGI page.

4.1.1. EGI Functions

The main page in this area contains brief descriptions of the envisioned EGI functions, and a link to the NGIs' opinions of a set of similar functions, collected in an EGEE survey of late 2006. This area is currently being reorganised to reflect the clustering of functions adopted by WP3 in its ongoing work, and is expected to grow substantially as WP3 begins to publish its results.

4.1.2. Use Cases

This area is dedicated to the Requirements and Use Cases presentation described more fully in D2.1.

4.1.3. European Projects

This area is dedicated to the analysis of the current European e-Infrastructure projects. EGI needs to provide a roadmap to a sustainable continuation of the activities covered by these projects, thus it is

essential that a thorough and systematic understanding of these projects be documented for the benefit of all involved.

Currently the project that is best known by the team is EGEE. Hence the related article is already quite informative. Other projects (e.g. DEISA, PRACE, OMII-Europe) will need the input from external experts. The latter will be invited to contribute as part of the WP3 activities.

4.1.4. Application Areas

One important measure of the success of grid usage is the variety of areas in which this usage occurs. In the EGEE survey of 2006, the number of distinct application areas in which at least one NGI is active was roughly 22 – a count which is probably an underestimate, since some areas were grouped together in analysing that survey. Many of these application areas were written in by the survey respondents (rather than being one of five preexisting choices), suggesting that there is much to learn about the variety of application areas in which grid technology is relevant. Thus the “Application Areas” pages will be essential to monitor the growth of e-Infrastructures in the general scientific community beyond the traditional areas of High Energy Physics, Bioinformatics, etc.

These pages will also record efforts within the private sector (e.g. Finance), which – from an EGI perspective – fit well as application areas.

It is furthermore necessary at this stage to provide a centralised, public census of all the applications currently served by the European e-Infrastructure(s), as the coming together of all the related efforts entails a level of information sharing that is becoming increasingly complex.

4.1.5. Legal Options

This area is set up for the work of WP4.

4.1.6. The Internal area

The EGI KB is intended to be a primarily public website. Nevertheless, as mentioned earlier, this tool must also be useful for the work of the EGI_DS team and the NGIs, and given the scope of this project it must enable efficient communication between members that by definition reside in different parts of Europe. Some of this communication is not necessarily aimed to the general public; thus there is a part of the EGI KB that is only accessible to registered members (both read and write). When a member logs in, they will be able to see and edit a reserved namespace called “Internal” (see MediaWiki Namespaces, <http://www.mediawiki.org/wiki/Namespaces>), which includes an entire area of the website that is not otherwise accessible.

The Internal area functions like any other part of the website: one simply puts the word “Internal:” before the name of the page he or she is creating, and the page will automatically be created as part of this restricted area. Links to these pages can be created anywhere, but when clicking on these links only registered users will be able to access them. For instance, the Main Page of the Internal area (**Internal:Main_Page**) can be reached via a link in the Main Page of the EGI KB (close to the bottom of this page), but a non-registered user will only see a “Restricted Area” page bearing the message “You are not allowed to view this page”.

4.2. THE NGI DESCRIPTIONS AREA

The individual NGIs each have a starter page, reachable either by searching by country or by clicking on the image map in Figure 1. For instance, clicking on Italy will bring up the page for the Italian Grid Infrastructure – IGI:

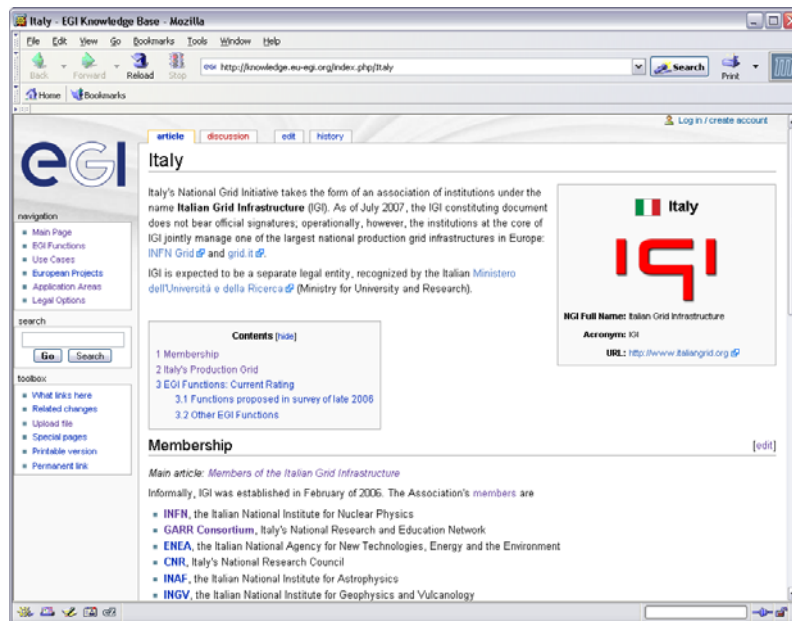


Figure 2: Example main page for a European NGI

Each NGI can create further pages, such as the one for the detailed description of the IGI members:

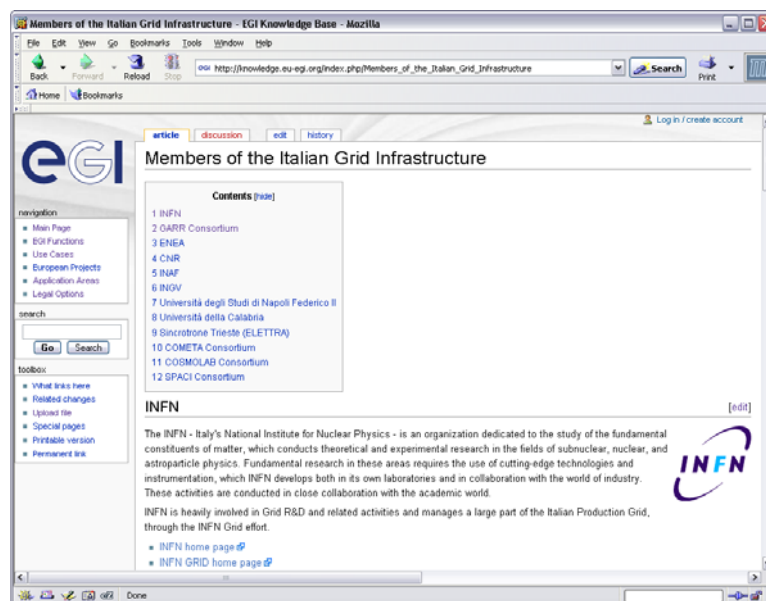


Figure 3: NGI-specific information – another page

As of March 2008, roughly half of the NGIs had made a contribution to their page. To encourage the process of contributing to the EGI KB, the EGI_DS project manager asked the NGI representatives to appoint a content manager for their NGI; in addition, the main EGI KB content manager contacted some of the NGI representatives individually (based on recent activity or expressed interest), thanking them for their contribution and/or asking them to add a contribution based on their expression of interest.



These communication activities produce good results if they are not abused, and if they are complemented by responsiveness to feedback on the part of the general content manager(s). One comment was that there are several places where the NGI representatives are asked to provide their information – indeed, in some cases the same information. This leads to the idea of integrating the EGI KB with the e-IRG Knowledge Base in a way that these sites can complement, rather than duplicate, each other.

4.3. INTEGRATION WITH THE E-IRG KNOWLEDGE BASE

The e-IRG Knowledge Base contains a wealth of information that is directly relevant to the work of the EGI_DS project and the NGIs. Rather than copying this information into the EGI KB, thus duplicating information and sources thereof, the EGI KB content manager and the e-IRG KB manager(s) have begun systematic contacts to find ways to share information across the two KBs. An initial plan was delineated in December 2007 with ambitions to “layer” the different kinds of information so that all would be available (in principle) in both KBs.

For instance, the e-IRG KB contains detailed technical information on computing resources in each country’s major computer centers, while the EGI KB is being populated with information on NGI composition (membership) and constituting documents. The former could then complete its information by automatically importing the organisational data provided by the latter into a dedicated “layer” created to host this information; the latter would in turn import the technical data into a dedicated “layer” of its own – e.g. in a specific page created for this purpose for each NGI.

Technical implementations are being studied as of early 2008. The choice of well-known tools in constructing the KBs makes it more likely than not that the layered approach will work.

Another issue to address is that of intellectual property (with respect to the content of each site). The appropriate text and disclaimer are now being studied and will be added in the near future.



5. ANNEXES

The EGI Knowledge Base: <http://knowledge.eu-egi.org>.