EGI-InSPIRE

European Grid Infrastructure – an Integrated Sustainable Pan-European Infrastructure for Researchers in Europe (EGI-InSPIRE)

EU DELIVERABLE: D2.3

Document identifier:	EGI-D2.3-v5
Date:	**18/04/2011**
Activity:	**NA2**
Lead Partner:	**EGI.eu**
Document Status:	**FINAL**
Dissemination Level:	**PUBLIC**
Document Link:	https://documents.egi.eu/document/201

Abstract

This white paper from the EGI-InSPIRE project gives detailed information about the European Grid Infrastructure, outlining the human, technical and infrastructure services that has been established and supported across Europe through the project. It provides a condensed overview of the technical activity supported by the EGI-InSPIRE that could be found in the Description of Work, but in a form that can be more easily referenced. Many of the topics briefly addressed in this paper are expanded on in other work within the project and these are referenced where appropriate. The detailed plans relating to longer-term issues and future challenges for the infrastructure in the areas of sustainability, virtualisation and expanding the user community are reviewed, but a detailed analysis of these topics are covered in work that goes on throughout the early part of the project but an overview of these issues is provided here.

Therefore, this paper provides a reference for individuals writing their own paper on EGI-InSPIRE and provides material that can be used within the project and its collaborators in other publications.

New users of the infrastructure should consult http://www.egi.eu to find information on how to get started in EGI.

I. COPYRIGHT NOTICE

II. DELIVERY SLIP

	Name	Partner/Activity	Date
From	Steven Newhouse	EGI.eu	22/11/2010
Reviewed by	**Moderator:** Catherine G. **Reviewers:** Isabel Campos Michal Turala	NA1/ NA2.2 NA3 EGI.eu EB	19/01/2011
Approved by	**AMB & PMB**		04/02/2011

III. DOCUMENT LOG

IV. APPLICATION AREA

This document is a formal deliverable for the European Commission, applicable to all members of the EGI-InSPIRE project, beneficiaries and Joint Research Unit members, as well as its collaborating projects.

V. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors. The procedures documented in the EGI-InSPIRE "Document Management Procedure" will be followed: https://wiki.egi.eu/wiki/Procedures

VI. TERMINOLOGY

A complete project glossary is provided at the following page: http://www.egi.eu/results/glossary/.

VII. PROJECT SUMMARY

To support science and innovation, a lasting operational model for e-Science is needed – both for coordinating the infrastructure and for delivering integrated services that cross national borders.

The EGI-InSPIRE project will support the transition from a project-based system to a sustainable pan-European e-Infrastructure, by supporting 'grids' of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI-InSPIRE will also be ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit user communities within the European Research Area.

EGI-InSPIRE will collect user requirements and provide support for the current and potential new user communities, for example within the ESFRI projects. Additional support will also be given to the current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences, as they move their critical services and tools from a centralised support model to one driven by their own individual communities.

The objectives of the project are:

1. The continued operation and expansion of today's production infrastructure by transitioning to a governance model and operational infrastructure that can be increasingly sustained outside of specific project funding.
2. The continued support of researchers within Europe and their international collaborators that are using the current production infrastructure.
3. The support for current heavy users of the infrastructure in earth science, astronomy and astrophysics, fusion, computational chemistry and materials science technology, life sciences and high energy physics as they move to sustainable support models for their own communities.
4. Interfaces that expand access to new user communities including new potential heavy users of the infrastructure from the ESFRI projects.
5. Mechanisms to integrate existing infrastructure providers in Europe and around the world into the production infrastructure, so as to provide transparent access to all authorised users.
6. Establish processes and procedures to allow the integration of new DCI technologies (e.g. clouds, volunteer desktop grids) and heterogeneous resources (e.g. HTC and HPC) into a seamless production infrastructure as they mature and demonstrate value to the EGI community.

The EGI community is a federation of independent national and community resource providers, whose resources support specific research communities and international collaborators both within Europe and worldwide. EGI.eu, coordinator of EGI-InSPIRE, brings together partner institutions established within the community to provide a set of essential human and technical services that enable secure integrated access to distributed resources on behalf of the community.

The production infrastructure supports Virtual Research Communities (VRCs) – structured international user communities – that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.

VIII. EXECUTIVE SUMMARY

This paper provides an overview of the EGI-InSPIRE project, a 4 year €72M project with a €25M contribution from the European Commission to support the transition towards a sustainable European wide production infrastructure. The project's 51 partners will establish and support human, technical and infrastructure services, described in this paper, that will evolve as a result of a 'virtuous cycle' of Discovery, Design and Delivery, driven by the needs of researchers and infrastructure providers to form the European Grid Infrastructure.

The European Grid Infrastructure marks a new approach to providing coordinated digital resources to European researchers. Many domains such as high energy physics, life sciences, environmental modeling, social sciences and humanities are facing the 'data deluge' challenge - how to process the deluge of large-scale data being generated by their communities. Successive capacity building projects such as the European Data Grid (EDG), Enabling Grids for e-Science (EGEE) and others have shown that transnational access to computing, storage and networking resources can be provisioned by federating resources into a production infrastructure. However, the project-based nature of these activities to date does not the stability required by multi-decade experiments and research. The European Grid Initiative Design Study (EGI_DS - http://web.eu-egi.eu/) explored options for building a sustainable high throughput computing infrastructure across Europe. The model that emerged is based on resources provided by National Grid Initiatives and brought together by a coordinating entity, EGI.eu.

EGI is a federation of national and domain specific resource providers coordinated by a dedicated organisation called EGI.eu, set up as a Dutch foundation (Stichting), and governed by the EGI Council. Technical activity is managed across EGI's three main areas - Operations, User Community and Technology by a number of groups representing various communities. These groups develop policies and procedures that follow the principles of openness, transparency and consensus and that can be developed internally or in collaboration with external partners.

EGI enables researchers within Virtual Research Communities (VRC) to collaborate, communicate and share resources across international boundaries by offering benefits such as the integration of community resources into EGI through user support and assistance, training and technology specialists and representation. Dissemination is also coordinated by the project with the aim to communicate the work of the EGI and its user communities within the project and worldwide. Audiences include new and existing user communities, the general public, journalists and policy makers, among others, who will be targeted using websites, wikis, printed materials, press releases, social media channels and booths at events.

Communities that engage with EGI make use of a suite of support services designed to help both experienced and new users exploit the infrastructure. Training is one of the key services provided for users to help guide the establishment of self-sustainable communities and EGI supports this through a training events list, a repository of training materials and a trainer profile database. The EGI Applications Database stores tailor-made computing tools for scientists from all fields as finished products, ready to be used on EGI. During the early stages of EGI-InSPIRE, specific support will be given to the Heavy User Communities (HUC), such as life sciences, high energy physics, astronomy and astrophysics and earth sciences to aid the transition of the critical capabilities that can be used by their communities and also by others.

EGI is a federation of resource services provided by a distributed and heterogeneous set of national or domain specific Resource Infrastructure Providers. Resource Infrastructure Providers are operational entities who coordinate their affiliated resource centres locally, and are themselves generally federations of resource centres or sites. Access to these resources is enabled through a baseline set of policies and operational procedures. The deployed service infrastructure includes local services provided by the resource providers, such as core middleware services, monitoring and accounting and certification, and global services offered by EGI.eu and its partners, such as VO middleware services, central operational tools and central gathering of monitoring and accounting statistics. EGI-InSPIRE supports the continued maintenance and development of the operational services through distributed product teams.

The EGI Technology Roadmap describes the components that deliver the architecture of the production infrastructure - within this the Universal Middleware Distribution Roadmap describes the evolution of components coming from technology providers within the EGI community. This roadmap is assessed regularly through the User Community Board and the Operations Management Board and new developments are prioritised by the Technical Coordination Board. EGI focuses on the operation of the infrastructure and relies on external technology providers, such as European or national projects, open source software or commercial providers, to provide technology ready for deployment into the production infrastructure. Before new technology is deployed, EGI.eu implements a set of Quality Assurance actions, providing guidance and assessment during three phases: defining and publishing Quality Criteria; proactive monitoring and assessment of test plans and results; and exposing software releases to production conditions on a few sites. EGI-InSPIRE will work with technology providers and virtual user communities to satisfy the requirements of users and resource providers within EGI.

EGI-InSPIRE is supported by the European Commission and its partners to make the transition towards a sustainable ecosystem of VRCs, resource and technology providers. Some of its elements are already funded through national and EC funding, some by a mixture of fees and service charges. EGI-InSPIRE will provide a framework for community based consultation on how these models will evolve. Virtualisation has led to widespread changes in commercial data centres, including the cloud computing business model. For EGI, virtualisation also offers the opportunity to deploy different software environments on demand, customised to the needs of the individual user communities. A further future challenge will be to expand the breadth of the user community by evolving EGI's services and lowering the complexity of their use. This will be assisted by employing a virtuous development cycle, including the Discovery phase, involving feedback from the VRCs, a Design phase comprising multiple development initiatives and finally a Delivery phase representing the provisioning, promotion and support of these services for the communities. The cycle continues as services are monitored and VRCs report back on future needs and goals.

In conclusion, EGI provides human, technical and infrastructure services through the federation of national and domain specific resource providers to researchers in Europe and their international collaborators. During its first four years, the EGI activities will be supported by the EGI-InSPIRE project, with support gradually decreasing during the project to allow the exploration and adoption of sustainable business models.

TABLE OF CONTENTS

1 INTRODUCTION

The European Grid Infrastructure (EGI) marks a new approach to providing coordinated digital resources to European researchers. Supporting the data analysis challenges facing many research disciplines, as a result of the 'data deluge', is becoming an essential capability to support research innovation across Europe. The data deluge that has been driving the recent development of Distributed Computing Infrastructures (DCIs), is frequently exemplified by the challenges facing the High Energy Physics community at CERN. However, many domains such as life science, environmental modelling, physical sciences and social science and humanities are facing similar challenges in processing the large-scale data being generated from within their own communities. Within Europe, some of these projects are part of the ESFRI (European Strategic Forum on Research Infrastructures) which designates facilities (either distributed or single site) of strategic interest to the European Research Area.

A critical question facing the research community is how to sustainably support the data analysis needs of these emerging Research Infrastructures. The capacity building activities that have taken place in the e-Infrastructure community through the support of the 7th (and earlier) European Commission Framework Programme have demonstrated that transnational access to computing, storage and networking resources can be securely provisioned to a multi-disciplinary user community through the federation of these resources into a production infrastructure. This has been notably demonstrated through successive GÉANT projects for networking, the European Data Grid (EDG) and successive Enabling Grids for E-SciencE (EGEE) projects for high throughput computing, as well as the DEISA and PRACE projects for high performance computing. However, all of these activities were project based – the funding for the e-Infrastructure that they provided was for a fixed short duration – and for researchers embarking on a multi-decade research experiment, such short-term funding makes the whole scale adoption of the e-Infrastructure unattractive.

The European Grid Initiative Design Study (EGI_DS - http://web.eu-egi.eu/) project was funded to explore the sustainability options for high throughput computing infrastructures across Europe. As a result of extensive community consultation, a model has emerged based around strong sustainable national grid initiatives that were brought together across Europe by a dedicated coordinating entity – EGI.eu. The transition from the structures that had been developed during the EGEE projects to a sustainable infrastructure is being supported by the EGI-InSPIRE project, a four-year €72M project with a €25M contribution from the European Commission, a cash contribution of €6M from the participants of EGI.eu, with the remaining effort being delivered as matching effort by the project partners. The EGI-InSPIRE project brings together 37 partners representing European National Grid Initiatives (NGIs), two European International Research Organisations (CERN and EMBL), eight unfunded partners in countries from the Asia Pacific region and is coordinated by EGI.eu based in Amsterdam.

The EGI-InSPIRE project establishes and supports services across Europe to provide Human Services (for coordination and community building), Technical Services (for supporting the interaction of user communities) and Infrastructure Services (for securely accessing resource hosted by different organisations). The evolution of these services is driven by the needs coming from the researchers and infrastructure providers within EGI and the organisations they collaborate with internationally. This evolution is driven by a virtuous cycle that includes the prioritisation of their requirements, the fulfilment of these requirements by external technology providers, the assessment of the new technology releases to ensure they meet the original requirements, and then the deployment of new technology into the European Grid Infrastructure.

2 SERVICE DRIVEN EUROPEAN GRID INFRASTRUCTURE

2.1 Accessing EGI

The European Grid Infrastructure in November 2010 connects over 300,000 cores in over 330 sites spread across over 50 counties around the world for over 10,000 users in over 20 active different communities. This document provides a technical description of the work support by the European Commission as part of the EGI-InSPIRE project and is not a source of documentation on accessing these resources. Some typical roles are identified below along with sections in this paper, alongside other sources, that might provide further relevant material:

- A new user trying to use the European Grid Infrastructure should contact their local National Grid Initiative or consult the User Support pages at http://www.egi.eu. Information on the current communities where there is specific support within EGI-InSPIRE can be found in Section 2.3.9.
- A new participant in the production infrastructure will find information in Section 2.4 and on the operations pages located on the EGI Wiki – http://wiki.egi.eu
- The interaction and governance between different elements in the EGI community through collaboration, policies and written agreements is described in Section 2.2.
- The technical services provided by the project to help user communities improve their use of the infrastructure by connecting with other communities, and providing access to general purpose tools, software and technical capabilities are found in Section 2.3.
- A technology provider that has software they would like to see deployed in EGI can find an overview of the assessment process in Section 2.4.4.

The EGI website (http://www.egi.eu) and wiki (http://wiki.egi.eu) will always contain the latest information on EGI's services and activities.

2.2 Human Services

2.2.1 Governance

EGI is a federation of nationally and domain specific resource providers that are coordinated by a new dedicated organisation based in Amsterdam – EGI.eu. The governance of EGI.eu, a Dutch Foundation (a Stichting) is entrusted to the EGI Council which draws its membership from the national or domain specific organisational resource providers that are participants in the foundation. Participation in the foundation is open to new members, subject to the approval of the Council, to those organisations willing to agree to its statutes and pay the fees agreed by the EGI Council. It is the EGI Council that decides the policies of the organisation, and therefore the community it represents, through a voting scheme that is weighted by the financial contribution each participant makes to the organisation, which is derived from the country's Gross Domestic Product (GDP)

Day to day, the activity of the EGI.eu organisation is managed by a Director who is advised by an Executive Board with six members and a chair elected from the EGI Council. Technical activity across EGI's three main areas – Operations, User Community and Technology – is managed through groups containing representatives from the relevant communities.

- The Operations Management Board (OMB) is chaired by the Chief Operations Officer and has representatives from the operational staff from each of the resource providers.

    - The Operational Tools Advisory Group (OTAG) deals with the detailed technical development of the tools being used within the operational infrastructure.

- The User Community Board (UCB) is chaired by the Chief Community Officer and has representatives from the Virtual Research Communities (VRCs) that are using the production infrastructure.
    - The User Services Advisory Group (USAG) deals with the detailed technical development of the services available to aid the development and integration of Virtual Research Communities.
- The Technical Coordination Board (TCB) is chaired by the Chief Technology Officer and has representatives from the technology providers that are contributing software to the Unified Middleware Distribution (UMD) Roadmap that forms part of the EGI Technology Roadmap.

The EGI-InSPIRE project brings together many of the participants in EGI.eu and provides support to the development of EGI during the first four years of its existence as it transitions towards sustainability. The programme of work within EGI-InSPIRE helps to establish EGI by committing the partners within the project to also work together within EGI by supporting the persistent structures within EGI that will remain after the project is complete.

2.2.2 Policy

Policy development within EGI is led by the EGI.eu Policy Development Team (PDT). The team aims to support the formulation and development of policies to be used by the European DCIs. The PDT also provides support to the boards and committees within EGI.eu that define policies and procedures for evolving the technical infrastructure, for its operation and for the access by the various VRCs. Policies are formal and mandatory statements or positions adopted by the EGI governance bodies for issues relevant to the community and link the long-term strategic cooperation of different parties with the day-to-day technical activity.

The policy development process rests on three fundamental principles. The first principle is *openness,* meaning that the policy development should be open to all stakeholders and actors within EGI community and follows an established participatory process of collaboration. Interested parties concerned with policy development are welcome to contact the relevant policy group and in agreement with them take part in their policy discussions. The second principle is *transparency*, that is policy discussions and policy papers should be archived and available through the EGI document repository (https://documents.egi.eu) to all interested parties and stakeholders. The third principle is *consensus,* meaning that whenever possible, policy document approval should be based on consensus.

Policies can be developed internally by policy groups or in collaboration with external partners. The internal policy groups are governed by Terms of References (ToR), i.e., documents describing their responsibilities, composition and operational procedures. So far, ToRs for the following groups have been defined*:* Technology Coordination Board (TCB) [R19], Operations Management Board (OMB) [R22], Operational Tools Advisory Group (OTAG) [R23], Operations Automation Team (OAT) [R18], User Community Board (UCB) [R21],User Services Advisory Group (USAG) [R16], Security Policy Group (SPG) [R20], Software Vulnerability Group (SVG) [R24] and Security Coordination Group (SCG) [R17].

All the groups have been created under the governance of EGI.eu since they are expected to last beyond the end of the EGI-InSPIRE project and become permanent groups of the EGI community.

The External Advisory Committee (EAC) [R25] has been established to advise the EGI-InSPIRE project in its activity and evolution.

The typical outcomes of the policy groups are policies or procedures to be adopted by the EGI actors. The PDT has defined a Policy Development Process (PDP) [R26], which governs the approval of policies and procedures produced by the EGI policy groups for the EGI community. The goal of the PDP is to ensure that all relevant parties, especially those affected by specific policies, have the opportunity to be part of policy discussions, review draft proposals and provide their input during the policy making process. Furthermore, the PDP provides consistency by making sure that all policies and procedures are treated in a similar fashion to prevent potential issues and conflicts and formalise a process of interaction between the participants in the policy development process.

External policies need to be established in consultation with all the relevant organisations within the EGI Community and that interact with it. A sustainable operation of EGI requires well-defined interactions with technology and infrastructure providers, as well as with the user communities through their VRCs to drive the evolution of our services. Interactions with these groups, and other policy bodies and projects within the community will be established through Memorandum of Understanding (MoUs), Operational Level Agreements (OLAs) and Service Level Agreements (SLAs) depending on the relationship. Furthermore, a mutual interaction with the funding agencies (mainly the European Commission) and international policy bodies are vital to align each other policy strategies.

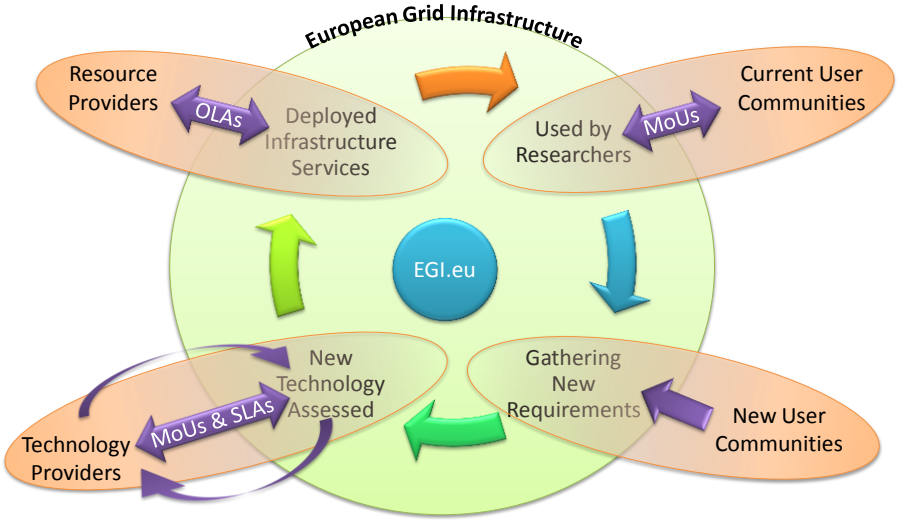

Figure 1: Virtuous Service Cycle linking the needs of current users of the infrastructure and new user communities to developing and deploying technologies from external technology providers.

Specifically, external technology providers (see section 2.4.4.2) are expected to sign a MoU to participate in the discussions taking place within the TCB. Those technology providers wishing to deliver software into UMD for deployment within EGI are expected to sign a SLA to assure the EGI community that the deployed software will be supported within the specified timescales. Likewise, VRCs will need to sign a MoU describing the scientific domains that they are active in and their representatives in interacting with the infrastructure. Resource Infrastructure providers from outside the participants and associated participants in EGI.eu will need to sign an Infrastructure Provider's MoU committing themselves to the OLA and operational polices that members of EGI.eu have

committed to. This ensures that EGI's users are assured of the same quality of infrastructure regardless of who provides the particular resource. Project MoUs will be used to collaborate with other EC or nationally funded projects where a mutually beneficial programme of work can be identified. The relationship between these different groups (resource providers, user communities, technology providers) with the NGIs and EIROs represented through EGI.eu is established in a virtuous service cycle (See Figure 1).

An early example of this collaborative activity has been with the other five European Commission-funded projects related to DCIs. Together, a collaborative DCI Roadmap has been developed that describes a vision of moving from the current production e-infrastructure in Europe to one based upon federated virtualised resources and records individual interactions between the six projects, in relation to this goal [R28]. Furthermore, the PDT has produced a Standards Roadmap [R27] which contains an overview of the relevant standards activity taking place in the EGI environment, both internally within the operational tools and through external software providers as described in the UMD Roadmap.

EGI is participating in the work of the e-Infrastructure Reflection Group (e-IRG), the European e-Infrastructures Forum (EEF) and the Open Grid Forum (OGF), and other high-level and international policy bodies. Additionally, EGI actively collaborates with dissemination and coordination projects such as e-Science Talk and SIENA. EGI is also following the latest development of EU strategic policies reflected in the papers Innovation Union and Digital Agenda for Europe [R29] and reports created by the EU high-level expert groups on clouds and scientific data [R30]. EGI perceives its role as a crucial one, since building a high-quality and reliable e-Infrastructure is a key component of a renewed European strategic concept for e-Science, e-Infrastructures, and innovation.

2.2.3 Virtual Research Communities

Large scale scientific research projects are no longer being conducted within a single research group, single institutions or even in a single country, but in VRCs that span national borders encompassing many different organisations with a need to share ICT resources.

EGI provides the e-Infrastructure that enables the researchers within a VRC to collaborate, communicate, share resources, access remote computers or equipment and produce results as effectively as if they, and the resources they require, were physically co-located. Becoming an identified VRC within EGI offers the following key benefits for researchers, scientists and the developers of distributed scientific applications:

1. VRCs can access computing, data storage and other types of resources made available by EGI stakeholders through open source middleware software solutions. VRC members can securely integrate their own community's resources into EGI in order to store, process and index large datasets.

2. The user support units of the NGIs and EGI.eu help VRC members during the routine usage of the systems and provide assistance to access and use the largest multi-disciplinary grid in the world. The EGI Helpdesk provides a single point of contact for all user support issues and is the mechanism for users to access NGI and EGI.eu support teams and its role is described in more detail in Section 2.4.3.2.

3. VRCs also have the ability to establish their own Virtual Organisations (VOs) – collections of hardware, software and human resources – configured to share capacities, to

collaborate with partners and to run data intensive simulations. The VOs can benefit from the resources provided by NGIs and other VRCs.

4. NGIs provide trainers and technology specialists for VRCs to support them during the integration and adaptation of their legacy applications and datasets to EGI. The combination of VRC's own training resources together with EGI's materials can provide comprehensive packages for VRC members in an efficient and timely manner.

5. VRCs can influence the evolution of EGI's services through representation in the User Community Board and the User Services Advisory Group. Based on requirements collected from its members, VRCs can advise EGI on its planning and operational priorities.

The VRC is therefore an organisational grouping that brings together transient VOs within a persistent and sustainable structure provided by EGI. A VRC needs to be a self-organising group, possibly using existing community structures as a basis of that representation, which collects and represents the interests of a focussed collection of researchers across a clear and well-defined field. Named contacts are agreed upon by the VRC to perform specific roles and these then form the communication channel between the VRC and EGI. Further details on establishing a VRC can be found at http://www.egi.eu/user-support/becoming_a_community/ which includes the MoU template that can be used by new communities to help them define the relationship between a user community, EGI and its NGI resource providers and user support teams.

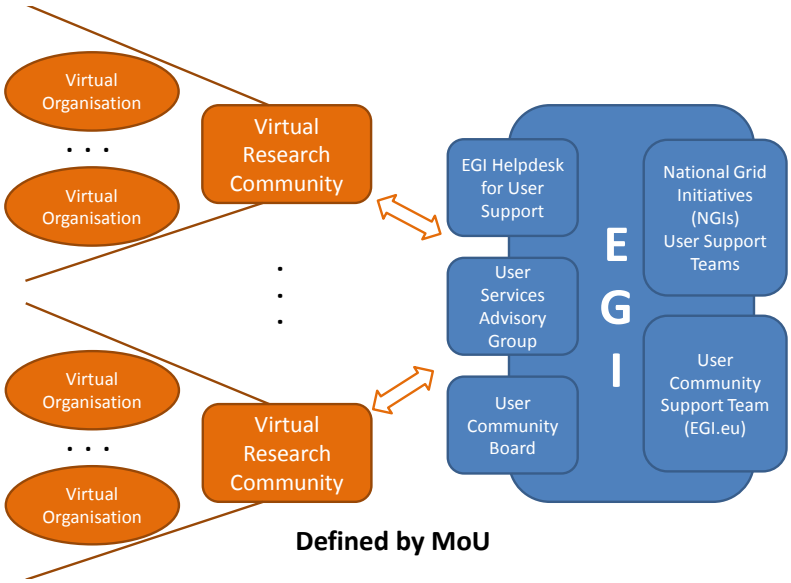

Figure 2: Relationship between Virtual Organisations, Virtual Research Communities and the User Support functions within EGI.

2.2.4 Dissemination

The dissemination activity is coordinated by EGI.eu on behalf of the European NGIs and projects, and other international partners with the aim to communicate the work of the EGI and its user communities both within the EGI-InSPIRE project and worldwide. The dissemination plan is targeted at EGI-InSPIRE partners contributing to Global and International dissemination tasks, but also to other EC-funded projects involved in dissemination such as e-ScienceTalk (http://www.e-sciencetalk.org/) and the SIENA Initiative (http://www.sienainitiative.eu).

The target audiences for the dissemination outputs are new and existing user communities, journalists, general public, grid research and standards communities, resource providers, collaborating projects, decision makers and governmental representatives. Means for dissemination will include the project website, wiki site, materials and publications, media and public relations, social media channels and attendance at events in order market EGI, and the activities of the EGI-InSPIRE project to new users. Some specific examples of the materials that will be produced in each of these categories are provided elsewhere [R1].

2.3 Technical Services

Communities that engage with EGI are able to make use of a suite of technical services provided to support both experienced and inexperienced users exploiting the infrastructure to further their research. These technical services may be web based (i.e. websites or portals), capabilities (e.g. support for parallel computing applications), standalone software tools/applications (e.g. workflow environments) or software services that clients connect to (e.g. GANGA).

Furthermore, formally engaging with EGI through the VRC model provides the community with a voice in the evolution of these support services in addition to the evolution of the infrastructure. This formal channel is defined by the interaction of the UCB, TCB and USAG committee structure described previously (section 2.2.2). The technical services described below are those selected on the basis of their proven usefulness within the context of the EGEE project series. As the needs of the user community evolves, the features and emphasis given to these services will change to meet the emerging needs of the user communities and the supported service set may expand to include other services.

During the early stages of the EGI-InSPIRE project specific support will be given to the Heavy User Communities to aid the transition of the critical capabilities that they need to either being part of the core e-Infrastructure available and used by all communities, or capabilities that are only relevant and are only used by a specific community.

2.3.1 Virtual Organisations

EGI users are organised into Virtual Organisations (VOs). A VO is a group of people (scientists, application developers, trainers, etc.) with similar interests and requirements, who are able to work collaboratively with other members and/or share resources (e.g. data, software, expertise, CPU, storage space) regardless of geographical location. One needs to be a member of a VO before he/she is allowed to access remote services, for example to run jobs or store files within EGI. EGI currently hosts more than 200 VOs for communities with interests as diverse as Earth Sciences, Computer Sciences and Mathematics, Fusion, Life Sciences or High-Energy Physics. The list of registered VOs can be obtained from the Operations Portal [R2].

2.3.2 Training

Training is one of the most important services provided for users to achieve better scientific performance with the grid and to guide the establishment of self-sustainable user communities. EGI provides a number of services aimed at supporting cooperation between trainers and users in different localities by connecting the groups through the activities that are established within the NGIs and scientific clusters. These services are available through the training website (http://training.egi.eu/) and include:

- Training events list which allows trainers to advertise their training events and to be made aware of other training events being run within the community. It also provides potential attendees with information about events within their local area and globally.
- Repository of training materials which provides access to a digital library which holds all forms of training related materials. This acts as a resource allowing trainers to share, find and re-purpose materials.
- Trainers profile database which holds information about trainers across the EGI area. Allowing trainers to contact each other and solicit expertise to help support their training activities.

At the start of the EGI-InSPIRE project these services were migrated from the EGEE website to the EGI.eu web domain, where they now evolve based on the feedback received from trainers and end users. Many of the NGIs hold regular training activities or are willing to do so on demand from interested user communities supported through their own national funds. The training services provided as part of EGI-InSPIRE help the coordination of these activities by linking events, trainers and training material.

2.3.3 Applications Database

The EGI Applications Database (AppDB - http://appdb.egi.eu/) stores tailor-made computing tools for scientists to use. The scope of the database embraces all scientific fields, from resources to simulate exotic excitation modes in physics, to applications for complex protein sequences analysis. The applications and reusable application developer tools filed in AppDB are finished products, ready to be used on the EGI. Storing pre-made applications and tools means that scientists do not have to spend research time developing their own software. The goal for AppDB is to inspire scientists less familiar with programming to use EGI and its resources due to the immediate availability of the software that they need to use. The database is also meant to avoid duplication of effort across the user community.

The AppDB has been online since the July 2010 and is the natural successor to EGEE's Application database. A recent feature of AppDB is the capability to store developers' and scientific users' profiles, so communities can even more easily find experts from specific domains. AppDB is publicly accessible and currently (November 2010) holds information about 269 applications and tools and the profiles of 494 application developers and end users.

2.3.4 Parallel Applications

Many scientific problems can be solved efficiently by running multiple subtasks in parallel. The Message Passing Interface (MPI) framework was developed to enable the scientific communities to exploit, in a simple, efficient and standard way, access to multiple CPU cores. It is the most widely used framework for parallel processing by the scientific communities. For example, the Astronomy and Astrophysics, Civil Engineering, Computational Chemistry, Earth Sciences, Fusion communities etc., have all developed a rich set of applications based around MPI libraries and toolkits.

MPI application development and usage in EGI is supported by each of the currently independent grid middlewares, however, the integration of the MPI into the infrastructure has been difficult. An effort in standardisation has played a crucial role in improving the situation. For instance, under the gLite middleware, each resource provider publishes information about which versions of the MPI are available. The resource provider can optionally publish more details about their local installation,

such as the compiler used by the MPI installation, if a shared file system is used, the network interconnect used in the cluster, etc., to allow the grid user to tailor or target their application at sites providing particular resources.

Recent computer architecture advances have enabled multiple CPU cores to be manufactured on the same physical chip. OpenMP is the dominant framework for exploiting multi-core CPUs. In addition, newer technologies, such as Graphics Processing Units (GPU), are now offering cheaper and faster alternatives to several hundred processing cores. Lessons learnt from making MPI available consistently within the production infrastructure can also be applied to these emerging resources.

EGI-InSPIRE provides improvements to the core services and software needed to support MPI and, in general, parallel processing. In particular, the commitment from the Computational Chemistry, Materials Science and Technology, and Fusion communities, to serve as 'early adopters' to new features alongside collecting requirements from other sources will ensure the needs and requirements of all user communities are met.

2.3.5 Ganga & DIANE

Ganga is an easy-to-use frontend for the configuration, execution, and management of computational tasks in a variety of distributed environments including Grid, Batch and Local resources that is partially supported by the EGI-InSPIRE project. It is an open and extensible system with specialised plugins to support VO-specific functionalities. Users may interact with Ganga through a choice of interfaces: a Command Line Interface (CLI) in Python, scripting or GUI. Ganga also provides integrated monitoring capabilities for jobs and tasks through a web-based monitoring applications and Dashboard monitoring services. Ganga supports a large variety of computing backends, including several middleware flavours, gLite, Condor, LSF, PBS, SGE systems and VO-specific workload management services such as PanDA and DIRAC for the HEP experiments ATLAS and LHCb. Future work includes further integration with monitoring systems and Dashboards, improving user support capabilities and streamlining experiment-specific features in the common code base for easier maintenance.

DIANE is a lightweight task processing framework that allows more efficient and robust execution of large numbers of computational tasks in unreliable and heterogeneous computing infrastructures. It exploits the late-binding method (also known as pilot jobs or worker agents) and provides an application-aware scheduler that may be extended by a system of plugins to support master/worker workloads such as task farms and bag of tasks. Plugins for Direct Acyclic Graph (DAG) and data-oriented workflows have been implemented as third-party contributions by a number of interested user communities. The framework also supports customised, application-specific processing methods and failure-management strategies. DIANE worker agents are often submitted using Ganga. This integration of the tools allows for a dynamic combination of resources across several distributed computing infrastructures, including High Performance Computing resources.

2.3.6 Workflow and Schedulers

Kepler is a workflow engine able to handle very complex and sophisticated scientific scenarios. It supports a wide number of scientific areas. This tool also provides the functionality to manage different computational environments hiding the differences among those environments to the final user. Thus, workflows involving applications running on different computational paradigms can be created and executed. GridWay is a workload manager that performs job execution management and resource brokering on a grid. This grid may consist of different computing platforms.

2.3.7 Dashboards

The Experiment Dashboard system provides monitoring of the WLCG infrastructure from the perspective of the LHC experiments and covers the complete range of their computing activities: job processing, data transfer and site commissioning. It provides transparent monitoring of the computing activities of the LHC VOs across several middleware platforms such as gLite, OSG and ARC. The applications provided by the Experiment Dashboard are widely used by the LHC VOs. Some of the Experiment Dashboard applications are generic and can be used outside the scope of the HEP community. The main direction of work within EGI-InSPIRE is to identify which components of the applications which are generic and can be shared by multiple VOs. Among those applications are the generic job monitoring, the site status board and the site usability.

2.3.8 GRelC

The Grid Related Catalogue (GRelC) service is a grid database management service aimed at providing access and management functionalities related to relational (PostgreSQL, MySQL, Oracle, DB2, SQLite, etc.) and non-relational (eXist, XIndice, XML flat files) data sources in a grid environment. It provides a uniform access interface to heterogeneous grid databases addressing:

- security through a web service interface that is integrated with EGI's authentication and authorisation mechanisms;
- interoperability with existing middleware (e.g. gLite & Globus);
- transparency by means of high level interfaces like the GRelC portal;
- efficiency providing advanced delivery mechanisms based on compression and streaming capabilities to efficiently transfer huge amount of data in a geographical environment.

The GRelC middleware is currently used in several grid research projects to support bioinformatics experiments on distributed databases as well as the metadata management related to Earth and Environmental Sciences applications in the Climate-G testbed. The network of GRelC services deployed within EGI will be enhanced to provide greater monitoring and control of the GRelC system infrastructure. Such a management framework will be available through the GRelC Portal by means of a new set of web pages that builds on the dashboard model providing a unified view (including charts, reports, tables) about the GRelC deployment, the status of the services, the available grid-databases, the supported VOs, etc.

2.3.9 Domain Specific Support

2.3.9.1 Life Sciences

The Life Sciences grid community covers a broad range of computerised health-related activities including bioinformatics, medical images analysis, drugs design and others. Medical data is by nature highly distributed (each biological laboratory or medical radiology centre is a source of data for the community) and can benefit tremendously from grid technologies to federate remote and heterogeneous data. In addition, the improved accuracy and capability of medical acquisition devices, leading to an exponential growth of biology and image databases, calls for the considerable data analysis capability that grids can deliver. The grid is also a vector for sharing expertise, through the publication of reference data analysis procedures and data sets that can be shared with a wide international community. These factors make the societal impact of the grid in Life Sciences very high. Consequently, this community has been identified early as a strong user community and has

been involved in the development of the grid infrastructure since the beginning. Life Sciences raise challenging issues for the grid designers, such as the high sensitivity of medical data and a large, scattered community, typically lacking expertise on distributed computing technologies [R3][R4][R5].

2.3.9.2 High Energy Physics

The High Energy Physics (HEP) community has pioneered the use of the grid, adopting it as the solution for the data processing and analysis of the data produced by the world's largest scientific machine, the Large Hadron Collider (LHC) at CERN. The first prolonged data-taking run of this machine started only shortly before the beginning of the EGI-InSPIRE project, in March 2010 and it remains a leading example of how a global infrastructure can support a science domain through the WLCG (Worldwide LHC Computing Grid) collaboration (http://lcg.web.cern.ch/lcg/public/). Despite many years of preparation, this first run has revealed some surprises both in terms of scale and usage patterns. The work supported by this activity aims to ensure that the exploitation of this world-class machine is successful, to help the grid adapt to changing needs and prepare a long-term sustainable model for support and operations. The deployment model used by HEP, as well as a number of the tools developed initially for this community (see Dashboards, Ganga and DIANE above) have been adopted by other disciplines and further sharing of technology, both within the HEP community as well as beyond, is considered an important issue for the HEP community within the project.

2.3.9.3 Astronomy and Astrophysics

Originally, Astronomy and Astrophysics (A&A) institutes usually acquired the necessary resources on local computing facilities and quite often also contracted the access to a pool of resources at supercomputing centres. However, some A&A applications grew in complexity and were considered very suitable to be run on grid infrastructures. At the moment the A&A HUC is devoted to the evaluation of different solutions for the "gridification" of a rich variety of applications. In particular the work is focused on:

- **Visualization tools:** VisIVO (http://visivo.oact.inaf.it/) for the visualization of data collected by different A&A projects
- **Databases and catalogues:** Currently the HUC is evaluating the possibility of adopting and tailoring different tools already in production (e.g. AMGA, GRelC, etc.).
- **Parametric applications:** Use DIANE (see Section 2.3.5) to study parametric space applications.

Additionally, one of the main goals is the accomplishment of a good level of interactivity among different technologies related to supercomputing, i.e. High Performance Computing and High Throughput Computing, Grid and Cloud.

2.3.9.4 Earth Science

Earth Science (ES) involves the observation, study and understanding of the behavior and variability of all of the components and characteristics of the Earth system. Moreover, it provides a strong scientific basis to face critical challenges related to new energy resources, secure water supplies, climate change, assessment of natural and industrial risks, etc. It is a community that makes intensive use of data with strict policy requirements, and needs to interface this data environment with the grid. The main requirement is to orchestrate data analysis and data modeling with efficient data transfers between data infrastructure or storage, the grid and HPC resources. Furthermore, the

community needs specific tools, such as Open Geospatial Consortium (OGC) components for geospatial applications and interfaces with web services highly developed in data centers. The need of intense computing is very important and related to the exploitation of the large volume of data, statistical studies (e.g. Monte Carlo or multi-scenarios approach) and collection of jobs. According to the application requirements, this community uses super computers as well as the Grid or other HPC resources.

The strategy of the ES Grid community, described in a roadmap for ES Grid published in an Earth Science informatics journal [R6], has consisted of propagating grid technology to all ES disciplines, setting up interactive collaboration among the members of the virtual ES grid community, and stimulating the interest of stakeholders on the political level. This strategy was applied through different European Grid projects [R3][R5][R7].

2.4 Infrastructure Services

EGI provides human and technical services to enable the coordinated operation and access to resources that are used by international user communities. These services are integrated with the operational procedures needed to properly run the production infrastructure during the day by day operations. The focus of these services is to provide a secure, reliable, integrated production infrastructure built from European resource infrastructure providers for user communities in the European Research Area. Where these user communities require integration with resources infrastructure providers outside of EGI, such as within Europe with providers of High Performance Computing through the PRACE (http://www.prace-project.eu/) project, or outside of Europe with Open Science Grid (http://www.opensciencegrid.org/) or TeraGrid (https://www.teragrid.org/) in North America, this activity will be supported.

2.4.1 Resource Services

EGI is a federation of national (NGI) or domain specific (EIRO) resource infrastructure providers that deliver integrated distributed services to EGI's user communities. Resource infrastructure providers (which may be based upon other DCI technologies such as desktop grids or clouds) are operational entities whose mandate is to coordinate locally their affiliated resource centres, and to peer with other infrastructure providers and centrally with EGI.eu. Seamless access to these resources within the provider is guaranteed through the deployment of UMD components of known integrability and a baseline set of policies and operational procedures. The resource providers are themselves generally federations of resource centres (also known as sites). These resource centres are the smallest operational entity in the EGI operations architecture. The deployed service infrastructure comprises:

- *local services* offered by the individual resource providers, such as core middleware services, monitoring and accounting, certification, grid oversight and support, etc.;
- *global services* offered by EGI.eu and its partners, such as VO middleware services, central operational tools, central gathering of monitoring and accounting statistics, coordination of middleware deployment, central Grid oversight, first and second-line support, etc.

2.4.2 Operational Services

EGI-InSPIRE supports the operation and development of the operational services needed to integrate the production infrastructure [R8] through distributed product teams. These teams work to evolve the capability of these tools to meet the needs of the resource providers (the NGIs & EIROs) that use

them. In particular, future developments will be driven by the evolution of the infrastructure that is expected to encompass additional resource infrastructures, such as desktop grids, cloud and high performance computing resources alongside the high throughput computing and storage resources that are already fully integrated into the infrastructure. The supported operational tools are:

- **Operations Portal**: (also known as the CIC Portal) a single access point to many operational information and operational tools. It is used extensively by all areas of the EGI community.
- **EGI Helpdesk (GGUS system)**: the main support access point for the project. Primary exploited by grid end-users and by support teams, but also by site managers who need support to solve middleware issues.
- **Grid Configuration Database (GOCDB)**: contains general information about the sites participating to the production infrastructure. Accessed across the whole project and by other tools or third party middleware in order to get grid topology.
- **Accounting Repository**: stores information about the usage of sites within the production infrastructure. Primary accessed by accounting system clients and by other tools (e.g. the Accounting Portal) in order to create statistical reports.
- **Accounting Portal**: graphical frontend for the Accounting Repository. Used by anybody interested in accounting information.
- **Service Availability Monitoring**: a monitoring framework for the resources that participate in the production infrastructure. Used by site managers to monitor their sites but also at a project level as data source to create availability/reliability statistics.
- **Metrics Portal**: collects a set of metrics from different resources to help measuring project performance and keep track of the project evolution displaying historical values of the metrics in a single place.

Many of the operational tools rely on messaging as a communication channel hence the project is also responsible for providing support to the integration of the tools with the message broker network run by EGI.

The development work is driven by requirements coming from bodies such as the User Services Advisory Group (USAG) [R9][R16] or the Operational Tools Advisory Group (OTAG) [R10]. When operational tools need to be integrated with the deployed middleware this is undertaken in collaboration with the external technology provider. If a new feature in the software component has to be integrated with the operational tools, then this integration will be approved through the TCB.

EGI-InSPIRE will complete the work started during the EGEE-III project to allow for a fully distributed or regionalised model of the operational tools. This work should be completed during the early stages of the project and currently:

- The operations portal team has already released a first version of the regional dashboard and have provided a schedule of future releases.
- The EGI Helpdesk provides a simplified regional helpdesk instance for NGIs. These instances are operated centrally but can be customised by the regions. Moreover NGIs can run their own systems connected to the EGI Helpdesk via standardised interfaces.
- Nagios based monitoring infrastructure and Messaging (SAM framework) is already completely regionalised [R11][R31].

2.4.3 Coordination Services

2.4.3.1 Grid Oversight

Ensuring that EGI provides a reliable integrated set of services requires pro-active monitoring and grid oversight at a local level to ensure quality of service to the user communities by a team of Regional Operators on Duty (ROD). This activity is complemented centrally by Central Operators on Duty (COD) who coordinate distributed ROD teams. CODs verify that operational issues are handled properly and locally by NGIs to ensure high quality services across the whole infrastructure. The work of the ROC and COD teams is driven by detailed procedures and manuals agreed within the community through the OMB. These procedures cover the duties and responsibilities of the different roles within the community encompassing the various tools used for operations, procedures for handling incidents and similar workflows and are described in manuals, wikis and training guides [R12].

One of the important duties of the COD team is to monitor EGI's availability and reliability statistics that are produced every month for all certified production sites] [R13]. For each monthly report, underperforming sites are requested through a GGUS ticket to explain the poor performance provided.

The personnel involved in grid oversight exploits the operational tools developed and deployed by the project and hence often interacts with developers to get support and to establish requirements for the development roadmap.

2.4.3.2 User and operational support

The EGI Helpdesk – a distributed support infrastructure with central coordination - is the system used to aid the process of supporting users and operations staff in their daily work in EGI. It is a distributed conglomeration of various ticket systems and other tools relying on a central integration platform which all other helpdesks (regional or topical) need to interface in order to allow structured communication between the various partners in the user support activity. Different support units are defined to cover different areas: operations and the related tools, baseline middleware, and application-level software. EGI provides support through a combination of local effort for first line support of users and operators, and central effort for ticket triaging and assignment, and specialised operational and middleware support.

The Ticket Process Management (TPM) is EGI's first line support: a team that receives incoming tickets, performs a triage and assigns the tickets further to expert support units. It works through the EGI Helpdesk to ensure that all support units are able to work on the ticket even if they do their work in a different tool that interfaces with the EGI Helpdesk. These 'second level' support units that are integrated into the EGI Helpdesk include:

- Application specific support units
- Community or Virtual Organisation specific support units
- Deployed Middleware Support Unit (DMSU), which includes component specific support teams, is used to investigate middleware related tickets that can be routed to external software providers (third-level support) [R32].
- Network Support Team coordinates the resolution of networking issues (e.g. performance) with the NGIs, the NRENs and DANTE. The LHCOPN has a separate support unit.

More information on EGI support system and the DMSU is available [R14].

2.4.3.3 Security procedures and incident handling

Grid sites are tightly interconnected, through the various middlewares, and through shared staff and users. This is a prerequisite for building a cohesive e-Infrastructure, but it also means that security incidents can easily affect different distributed sites. Incident response must be consequently proportionally strong and decisive.

The EGI Computer Security and Incident Response Team (CSIRT) covers all aspects of operational security aimed at achieving a secure infrastructure within EGI and relies on both site and NGI security contacts. The EGI CSIRT ensures both the coordination with peer grids and with the NGIs and NREN CSIRTs. The EGI CSIRT acts as a forum to combine efforts and resources from the NGIs in different areas, including grid security monitoring, security training and dissemination, and improvements in responses to incidents (e.g. security drills) [R15].

2.4.3.4 Accounting

Information about usage of EGI resources needs to be constantly gathered in order to assess the utilization uptake, the workload distribution among different user communities, the penetration of usage for different scientific disciplines, and to inform future business models for the provision of the resources. Exchange of usage records relies on the availability of a common communication infrastructure (increasingly based on messaging) and on standard usage record schema.

2.4.4 Service Evolution

2.4.4.1 Roadmap

The EGI Technology Roadmap describes the components that deliver the architecture of the production infrastructure as a whole. The UMD Roadmap is a sub-set that describes the evolution of components coming from technology providers within the EGI Community. Together, these roadmaps provide information about the current set of functionality provided to EGIs users, and identify which needs of the user communities (both end-users and operations staff) will be addressed, implemented and delivered in the future. Hence the roadmap describes how EGI will evolve over time [R33].

Various activities feed into the periodic revisions of these roadmaps. Based on the current set of components and functionality provided within the production infrastructure, the UCB and OMB regularly collect and assess requirements from associated VRCs and Infrastructure Providers, respectively. Requirements identifying new or changed needs are presented to the TCB. The TCB prioritises these requirements and identifies technology providers that commit to supply software components implementing those needs at a published timescale. Those recurring prioritisation actions result in regular updates of the Roadmap.

2.4.4.2 Technology Providers

EGI focuses on the operation of the infrastructure and relies on external technology providers (i.e. European or national projects, open-source software or commercial providers, etc.) to provide technology ready for deployment into the production infrastructure. It is this ecosystem of stakeholders (infrastructure operations, resource providers, technology providers, end-users, etc.) that moves the community towards long-term sustainability. Identifying consumers and providers in

a service-oriented model through defined communication boundaries allows any professional specialists to contribute functionalities to the EGI as either a Technology or Resource Provider.

The clear communication of requirements, and the terms and conditions of the technical interaction between technology providers and EGI, will help facilitate a long-term supply of high quality software components that are in alignment with the expressed demand of the users and the resource providers within EGI. The relationship that EGI.eu has with external technology providers is captured in a MoU which identifies contact points and a shared workplan. Technology providers that are delivering software to EGI are expected to sign a SLA describing the levels of support given to individual components including response times to resolve issues of different priority.

2.4.4.3 Technology Assessment and Deployment

To deliver sustainable quality with minimal disruption, EGI.eu implements a comprehensive set of Quality Assurance actions through experts in the community. Accompanying the technical software development and delivery process carried out by the technology providers, EGI.eu provides guidance and assessment during three important phases:

1. Quality Criteria that must be satisfied by the delivered software components are defined and published so that any provider can contribute components to the ecosystem. Generic and Component Specific Quality Criteria provide orientation to the technology providers to proactively enforce their own Quality Assurance processes along the published sets of Quality Criteria.
2. Pro-active monitoring and assessment of test plans and results published by the technology provider, allows EGI to verify all Quality Criteria applicable to software components delivered in a given release to EGI, ensuring constant delivery of high-quality software components to the production infrastructure.
3. Before any software release is provided for general deployment in the production infrastructure, it is exposed to production conditions through its deployment by a few sites in the production infrastructure. Through this limited software rollout, issues that may reveal themselves only in production conditions can be provoked and assessed before wider-scale deployment. Final acceptance reports determine whether the release under assessment is considered fit for general availability, or is rejected before it disrupts the production infrastructure as a whole.

3 FUTURE CHALLENGES

3.1 Sustainability

Through the EGI-InSPIRE project, the EGI community has the support of the European Commission and its partners in the project, to make the transition towards a sustainable ecosystem of resource providers, technology providers, virtual research communities coordinated on behalf of that community by EGI.eu. Some elements of this ecosystem will continue to be funded directly through national or European level funding mechanisms. Other elements may be supported by a mixture of fees and service charges.

This ecosystem was defined during the EGI_DS project and the implementation of which started during the EGEE-III project and will complete during the early years of the EGI-InSPIRE project. As the interaction between the different actors in the ecosystem stabilises, the components that need to be sustained long-term outside of short-term project funding and the components amenable to project funding will become clearer. Part of the work of EGI-InSPIRE project is to continue the community-based consultation to drive these discussions, the next phase of which is described in [R34].

3.2 Virtualisation

The commoditisation of hardware, software and networking over the last decade has fuelled the establishment and expansion of the infrastructure that was inherited by EGI. The next few years will see the impact of commoditised virtualisation within the provision of resources for the European Research Area. The changes that have been seen in commercial data centres and have led to the cloud computing business model, will inevitably impact the way data centres in the research community are provisioned.

The impact of adopting virtualisation locally and across the whole production infrastructure would allow the rapid and flexible deployment of different software environments, customised to the needs of individual user communities, with minimal overhead to the individual resource centre and on demand from the end-user. The technical evolution of EGI towards such a model is still under discussion and a more detailed analysis of the issues is provided elsewhere [R35].

3.3 Expanding the User Community

Expanding the breadth of EGI's user community by evolving its services and lowering the complexity of its use are key objectives supported through the EGI-InSPIRE project. The VRC model allows EGI to identify the communication points within each community to collect feedback and new requirements on the offered services in order to evolve them through a virtuous cycle of Discover, Design and Deliver.

The *Discovery* phase of the cycle involves supporting multiple mechanisms whereby the communities can find their voice and be heard by those developing and offering services for them. The *Design* phase comprises multiple development initiatives taking place both within EGI and amongst partners and parallel projects in a coordinated manner. Finally, the *Delivery* phase represents the provisioning, promotion and support of these services for use by the communities. Then the cycle continues as services are monitored and reviews and the VRCs report back on future needs and goals. This cycle is supported and championed through the EGI Technical and User Forums, various themed workshops and other events and activities frequently in partnership with other projects and initiatives.

4 CONCLUSIONS

EGI provides human, technical and infrastructure services through the federation of national and domain specific resource providers to researchers in Europe and their international collaborators. During the first four years, while this new ecosystem of resource and technology providers, virtual research communities and infrastructure coordination is being established, these activities will be supported by the EGI-InSPIRE project. Sustainability of these activities after the end of the EGI-InSPIRE project is a key concern that will be explored throughout the project, identifying different business models that can be accepted by the community and activities that they can support.

5 REFERENCES

R1	D2.2 Dissemination Plan https://documents.egi.eu/document/56
R2	List of EGI Virtual Organisations on the Operations Portal http://cic.gridops.org/index.php?section=home&page=volist
R3	MS602 The HUC software roadmap https://documents.egi.eu/document/230
R4	MS604 Services for Life Sciences https://documents.egi.eu/document/236
R5	D6.1 Capabilities offered by the HUCs https://documents.egi.eu/document/154
R6	A roadmap for a dedicated Earth Science Grid platform, R. Cossu, M. Petitdidier, J.L. Linford, V. Badoux, L. Fusco, B. Gotab, L. Hluchy, G. Lecca, F. Murgia, C. Plevier, P. Renard, H. Schwichtenberg, W. Som de Cerff, V. tran, G. Vetois. Earth Science Informatics, 3, 3, 135-148, 2010. DOI: 10.1007/s12145-010-0045-4, 2010
R7	MS603 Services for High Energy Physics https://documents.egi.eu/document/147
R8	MS702 Establishing the Operational Tool Product Teams https://documents.egi.eu/document/52
R9	https://wiki.egi.eu/wiki/USAG
R10	https://wiki.egi.eu/wiki/OTAG
R11	MS703 Operational Tools Regionalisation Work Plan https://documents.egi.eu/document/107
R12	Operational Procedures https://documents.egi.eu/document/15 https://wiki.egi.eu/wiki/Operations:Manuals https://wiki.egi.eu/wiki/Operations:Best_Practices https://edms.cern.ch/document/1015741
R13	https://wiki.egi.eu/wiki/Availability_and_reliability_monthly_statistics
R14	MS403 EGI Helpdesk and the NGI Support Unit https://documents.egi.eu/document/99
R15	https://wiki.egi.eu/wiki/EGI_CSIRT:Main_Page https://documents.egi.eu/document/47
R16	User Services Advisory Group (USAG) - Terms of Reference https://documents.egi.eu/document/121
R17	Security Coordination Group (SCG) - Terms of Reference https://documents.egi.eu/document/119

R18	Operations Automation Team (OAT) – Terms of Reference https://documents.egi.eu/document/123
R19	Technology Collaboration Board (TCB) - Terms of Reference https://documents.egi.eu/document/109
R20	Security Policy Group (SPG) - Terms of Reference https://documents.egi.eu/document/64
R21	User Community Board (UCB) - Terms of Reference https://documents.egi.eu/document/120
R22	Operations Management Board (OMB) - Terms of Reference https://documents.egi.eu/document/11
R23	Operational Tools Advisory Group (OTAG) - Terms of Reference https://documents.egi.eu/document/103
R24	Software Vulnerability Group (SVG) - Terms of Reference https://documents.egi.eu/document/108
R25	External Advisory Committee (EAC) – Terms of Reference https://documents.egi.eu/document/170
R26	Policy Development Process (PDP) https://documents.egi.eu/document/169
R27	D2.5 Standards Roadmap https://documents.egi.eu/document/206
R28	D2.4 DCI Collaborative Roadmap https://documents.egi.eu/document/207
R29	Digital Agenda for Europe http://ec.europa.eu/information_society/digital-agenda/index_en.htm
R30	Final Report of the High Level Expert Group on Scientific Data – Riding the wave cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf
R31	MS401 Operational Tools Regionalisation Status https://documents.egi.eu/document/48
R32	MS502 DMSU Operations Procedures https://documents.egi.eu/document/69
R33	D6.1 Unified Middleware Distribution Roadmap https://documents.egi.eu/document/100
R34	D2.7 EGI Sustainability Plan https://documents.egi.eu/document/313
R35	D2.6 Integration of Clouds and Virtualisation into the European production infrastructure https://documents.egi.eu/document/258