EGI Position Paper on the ERA Framework

EGI.eu welcomes the opportunity to participate in the public consultation on the ERA Framework: Areas of untapped potential for the development of the European Research Area (ERA).[1] This position paper contributes to the information gathering aimed by the online questionnaire addressing the topic of e-Infrastructures and the mechanisms by which they can promote EU Research and Innovation.[2]

For the last decade, the European Commission and national funding agencies have invested heavily in building Europe's e-Infrastructures. In the area of Distributed Computing Infrastructures (DCIs), early experiments took place in the European Data Grid (EDG) project, while the Enabling Grids for E-SciencE (EGEE) series of projects developed a production quality infrastructure. As a result, the European Research Area (ERA) is able to benefit from a pan-European e-Infrastructure that is available to communities ranging across astronomy and astrophysics, life sciences, computational chemistry, materials science, fusion, earth science and high energy physics. This European-wide e-Infrastructure has transitioned to a federation of national e-Infrastructures thanks to the EC funded EGI-InSPIRE project (European Grid Infrastructure: an Integrated Sustainable Pan-European Infrastructure for Researchers in Europe). Since May 2010, the European Grid Infrastructure (EGI) has been coordinated by a new dedicated organisation, EGI.eu[3], based in Amsterdam. EGI.eu plays this coordinating role on behalf of its stakeholders – the national and domain specific resource providers.

Today, EGI integrates almost 260,000 cores and more than 122 PB disk and 127 PB tape of storage space located at 350 sites in more than 40 national resource providers. The total of integrated and peer resources includes more than 400 sites and 350,000 cores. This production quality platform provides a flexible solution for many distributed storage and computing use cases. EGI currently supports over 18,000 researchers in their intensive data analysis needs in almost every e-Science discipline, and can extend this support to other e-Science communities and beyond to areas such as e-Government. EGI's innovative technology and procedures address many of the issues identified in the Digital Agenda for Europe. Through its European wide federation of national resource providers it is ideally placed to provide an e-Infrastructure (grids of computing, storage and cloud resources) for the general benefit of society.

Researchers

From the EGI perspective, researchers need many services to support the whole research lifecycle. These may range from the non-technical (e.g. bibliographic services, repository services, publishing services, collaboration services) to the technical (e.g. authentication services, data analysis services, workflow services, information services, data movement services) and the social (e.g. collaboration services, reputation services). These services need to scale either as individual instances or through interoperation with other instances across research communities of different sizes and operation. Thus, EGI services are scaled for individual researchers, research groups and Virtual Research Communities (VRCs).

The individual researchers are interested in "me-science". The use of e-Science services (supported by e-Infrastructure) is just one of many competitive advantages that they will employ to get their results published first. Integrating public or community data sources with their own local data sources and being able to run their own analysis may require that they interact directly with

---

[1] http://ec.europa.eu/research/consultations/era/consultation_en.htm
[2] All key recommendations are listed in Table 1.
[3] http://www.egi.eu/

the e-Infrastructure available to them or indirectly by benefiting from existing easy to use services that utilise the underlying e-Infrastructure as they required. EGI is aware that there are 1.85 million researchers (Head Count (HC)) employed in the European public sector (higher education and government sector)[4]. In order to maximise the performance of researchers in the ERA in terms of the quality, relevance and impact of their outputs, improved use of digital resources such as those provided by EGI and other e-Infrastructures is needed. This is the reason the European Commission (EC) should determine how many researchers need to use digital resources in their daily research activities (e.g. Engineering and Technology, Medical and Health Sciences) and how many could benefit from improved access to digital resources. This would be crucial to discover researchers' needs for e-Infrastructures. E-infrastructures do not have the resources to do this thorough investigation without EC support; therefore EC assistance is required in order to identify the genuine e-Science needs of researchers.

Recommendation 1: EC should provide information (e.g. through Eurostat) on how many researchers need access to distributed digital resources as part of their daily professional activities and how many researchers could benefit from improved access to digital resources.

Recommendation 2: EC should support the establishment of social networking tools to help researchers to connect with each other and advertise/discover research groups. This could be bound to services like OpenAIRE where publications, scientific data and research groups are discoverable and such collaborations used to govern access to e-Infrastructure services.

For many research communities, the core e-Infrastructure services are accessed through a domain specific e-Science environment. Therefore it is critical that the development of these environments continues to receive support (e.g. the recent e-Science environments call[5]). Finally, support also needs to be given to new research communities with new fundamental e-Infrastructure requirements to develop solutions in close collaboration with the e-Infrastructure provider. This model was successfully used to develop the e-Infrastructures we have today, and while the user communities may change, the model is vital to foster new developments. EC support will also be needed to help researchers run their own community's services and to develop their own software services in order to integrate access to the e-Infrastructure used by their community within a simple interface.

Recommendation 3: The EC should support applied science communities with funding to develop domain specific platforms to be deployed on e-Infrastructures, thereby offering new innovative services to a community's researchers.

Cross-border operation of research actors

When it comes to transnational collaboration of different national research actors, the National Grid Initiatives (NGIs)[6] and EIROs through forming the partnership for EGI are an example of a success story. Through EGI, the governance of grid and soon federated cloud computing resources is being achieved which is providing across Europe a coordinated, interoperable and harmonised approach to related policies, programmes and practices. This approach favours the deployment of measures at the most appropriate level in accordance with subsidiarity.

Through successful cross-border operation of NGIs, EGI has managed to:

---

[4] Eurostat
[5] http://cordis.europa.eu/fp7/ict/e-infrastructure/events-20100611_en.html
[6] National Grid Initiatives or Infrastructures (NGIs) are organisations set up by individual countries to coordinate the computing and storage resources that they provide to the European e-Infrastructure (EGI) to meet the needs of their local user communities to collaborate internationally. NGIs are EGI's main stakeholders, together with CERN. Each NGI contributes a number of sites to the grid infrastructure. NGIs are responsible for the maintenance and running of the sites they operate.

- Enhance trust. EGI.eu is leading the community effort to develop the policies needed to provide guidance, improve decision-making, clarify roles and responsibilities, and manage risks. Policies are developed through an open consultative process involving all stakeholders and are adopted by consensus to ensure transparency and foster trust.
- Guarantee integration. The resources of the EGI federation span the whole of Europe, not just the countries with developed economies, which are most likely to invest in their own e-infrastructures. EGI is a pan-European e-Infrastructure where every country is welcome to contribute resources and reap the benefits of an integrated, transnational effort.
- Promote interoperability. EGI is a federation of more than 35 countries, crossing culture and linguistic divides. Like other transnational organisations, it runs the risk of becoming swamped under the weight of different ways of doing things. EGI.eu ensures that everyone speaks the same language. This ensures a reliable service to the researchers who do not have to worry about who is providing the computer power that they are using.

Recommendation 4: EC should disseminate the EGI success story to other national and European RIs in order to demonstrate a possible way of achieving successful cross-border collaborative operation of e-resources for researchers in other disciplines.

Furthermore, greater integration between e-Infrastructures and broader support of different usage models is a request frequently heard from the user community, and increasingly by national stakeholders. EGI provides the means to federate nationally provided resources – currently focussed on HPC and HTC resources – but this is being extended to virtualised resources as there is a clear user demand and commitment from the stakeholders. Exposing these capabilities to a user base beyond e-Science, into e-Government and other domains is feasible with appropriate European and national support. The broader the supported user base the greater the need will become to tackle the potential integration of petascale computing, networking and data resources into the same integrated governance and policy structures.

Recommendation 5: EC should support the integration of European e-Infrastructures across different resource types (grids of HTC, HPC, petascale, virtualised, data and networking resources) in order to provide a coherent technical and policy structure. If achieved, this integration would enable new capabilities for new innovation within Europe (e.g. multiscale modelling)

Research Infrastructures

The key issues identified in the ERA Framework consultation document for Research infrastructures (RIs) are:
- Exponential growth of research data.
- Sub-optimal exploitation of the scientific, innovation and societal potential of existing RIs (priority setting, management and funding issues).
- Sustainability of the competitive edge of RIs as well as the development of new RIs of pan-European interest.[7]

The exponential growth of raw research data generated by experiments and computing simulations can exceed 10 Petabytes a year (and is growing almost exponentially each year), which needs to be stored, shared and analysed. This data volume is several times bigger than the space needed to store 10 billion photos on Facebook. Investment in the underlying physical infrastructure frequently undertaken through national funding bodies – networking, computing and storage – is needed to meet this scale. The need for the EGI will increase in the future as the emerging paradigm of data-intensive science becomes the norm and as more users and communities produce more Petabytes of data, delivered annually from experiments and complex

---

[7] http://ec.europa.eu/research/consultations/era/consultation_era.pdf

scientific simulations.[8] The Large Hadron Collider based at CERN is one example, generating more than a million computing jobs every day.

The Council of the European Union stated that: "The multiplication of initiatives and the fragmentation of efforts are slowing down the realisation of concrete results. Complementarity, prioritisation and efficiency should be at the core of European research policy".[9] The EGI federation of resource providers, National Grid Initiatives, research organisations and end-users has been reducing fragmentation and promoting integration for a decade through its predecessor projects. Together, the community has built an integrated pan-European research infrastructure to enhance research and innovation in Europe for the benefit of all and to reduce such fragmentation.

EGI.eu coordinates across Europe the operational and technical services provided by the EGI community to the community and operates the largest federated "grid" infrastructure in the world for multi-disciplinary research. Centralised services ensure cohesion between the resources provided in Europe and elsewhere. This allows resource providers to be more cost-efficient and concentrate on improving the services delivered to research communities.

Recommendation 6: Establish a specific dedicated technical working group under the ERA framework to help remove the potential barriers and bottlenecks that hinder collaboration between RIs and e-Infrastructures. To disseminate the working group's yearly activities, the EC should facilitate an annual forum between RIs and e-Infrastructures that would review the activities of the working group and discuss and adopt its proposals.

Recommendation 7: One of the expected outputs from ESFRI projects in the construction and implementation phases should be to formally evaluate the services, capacity and data needed from e-Infrastructures for their research needs and to report this information to the e-Infrastructures.

Recommendation 8: The EC should build an ERA Compendium/Knowledge base with all European RI, ESFRI projects, national and European research groups. A compendium/knowledge base should also describe the assets (e.g. people, services) in the RIs that are part of ERA. Having this structured collection of information would help the research communities to build collaboration and e-Infrastructures and other stakeholders to support their needs. The EC should provide a tool that would enable RIs to update the information anytime.

The recent global financial issues have impacted the national funding streams available to e-Infrastructure providers and the support given to their user communities. NGIs, as key stakeholders of EGI, are concerned about their own sustainability, as many of them rely on increasingly shorter-term funding. It is critical that the EC retains and reinforces its commitment to a federated European e-Infrastructure composed of national resource providers to support the ERA. Co-funding of strategic investments in e-Infrastructures, by the EC and Members States, can be very effective in supporting their sustainability. EC financial tools, Horizon 2020 and Structural funds should be among the corner stones of future sustainability, and this needs to be communicated clearly. The future sustainability of e-Infrastructures lies in the diversity of funding (e.g. SF funds, EC project funding, service charging). In the EGI context, sustainability issues can be solved by defining funding sources both on an EU and national level for central coordination, national resource providers, technology providers and VRCs.

---

[8] Increasing digitisation of research is literally leading to a data deluge as more and more disciplines adopt e-Science – these communities need to establish their own support and governance structures to interact in an integrated manner with e-infrastructure providers (through VRC model or establishing research group).
[9] http://register.consilium.europa.eu/pdf/en/10/st16/st16191.en10.pdf

Visible progress can be observed in the area of RIs where a clear roadmap, clear ownership by Member States and EU legislation has led to the establishment of new pan-European RIs despite scarce resources. There is a need to develop a similar roadmap for e-Infrastructures.

> **Recommendation 9:** The EC, together with e-Infrastructures, should develop the long-term roadmap for integrated e-Infrastructures aimed at a sustainable operation and evolution of e-Infrastructure services.

Engaging with Industry

Enhancing the dialogue with industry (especially SMEs) in order to foster knowledge and technology transfer is needed within the EGI ecosystem, both as a supplier and as user. Commercial adoption of EGI assets falls into several categories: knowledge exchange between research and commercial practices and terminology; joint collaborations for software development and/or provision; proof of concept activities; support for SMEs and start-ups. Activity in this area is hindered by the real or perceived legislative barriers in allowing commercial organisations to use publicly funded resources. In addition, there is a danger under anti-competiveness legislation of publicly funded e-Infrastructures providing services at lower cost due to their public subsidy than what is available from competing commercial counterparts.

> **Recommendation 10:** The EC should tackle the legislative uncertainty around commercial use and collaboration with e-Infrastructures.

EGI is continuously evolving to take advantage of new technologies in order to meet the emerging requirements of its users. Cloud computing and virtualised resources have attracted a lot of interest, thanks to the flexibility they offer to their users. EGI is committed to adding cloud computing through virtualised resources to the service portfolio offered to European researchers. At the moment, cloud technology is still in its development phase, driven by the commercial companies that have turned migration and storage of large volumes of data into a profitable business. EGI.eu wants to lead the way to integrate federated virtualised resources into the European Grid Infrastructure to deliver a pan-Europe cloud to support the ERA. This is important because commercial cloud providers are driven by their own different set of goals that may or may not converge with the interests of national governments and research institutions. Commercial providers are naturally inclined to prioritise the requirements that make business sense and that can be turned into a profit. Europe needs cloud providers, probably publicly funded, that are able to cater to the specific requirements of the research community that cannot be commercially provided and to make sure that economic obstacles do not stop innovation.

The EGI community and EGI.eu are perfectly placed to lead the development of a federated European cloud by using virtualised resource providers from the public and commercial sectors by building on decade-long experience of managing the evolution of distributed computing and training a generation of researchers in its use. Only a federated solution guarantees that no country in Europe is left behind and that user communities in the public sector are not locked into a single commercial provider. A federated model also helps drives interaction and competition between different providers, reduces fragmentation by defining a common base that others can innovate upon and promotes the importance of standards. It is therefore critical to achieve the ambitious goal of integrating virtualised resources into an European Cloud that can become a catalyst for innovation in Europe by bringing the ERA online.

> **Recommendation 11:** The EC should support EGI in leading the development of a European cloud built from federated virtualised resources from public and commercial providers and not to duplicate and fragment national efforts.

International dimension of ERA

The EGI partnership includes not only EU members, but also organisations from non-EU countries such as Turkey, Croatia, Serbia, Russia and Armenia. Furthermore, EGI has developed long-term collaborations through Memorandum of Understandings with organisations from South Africa, Latin America and Ukraine to integrate the respective resource providers. EGI has also active collaborations with the United States and the Asia-Pacific region. This wide collaboration network is a strong driver in enabling research groups to connect among each other.

Recommendation 12: The EC should strengthen the relationships between the funding agencies of key foreign countries to develop common policies for better integration of the respective e-Infrastructures. This would have the beneficial side effect of enabling the respective research groups to discover each other and collaborate. Furthermore, researchers would have access to more scientific data and computing capacity as well as to knowledge and expertise.

Global standards are critical for e-Infrastructures and Europe must invest in their political and technical development by following a collaborative approach. Standardised interfaces to European services are essential in order to reduce barriers to adoption, provide borderless services and provide platform for further integration to take place. However they not sufficient on their own if they are done in isolation. European e-Infrastructures work in a global context and their services need to interoperate with other providers so that an integrated set of services are provided for end-users. Projects such as OGF-Europe[10] and the SIENA Initiative[11] provide valuable community building work but lack the technical resources to underpin the standards discussions with substantial technical activity.

Recommendation 13: The EC should strengthen the relationships with the funding agencies of key foreign countries to develop policies to stimulate the definition of joint standardisation roadmaps encompassing the full standards development life-cycle. This would allow the e-Infrastructure community to remove the remaining interoperability barriers and to provide a seamless set of interoperable services that can underpin the Digital Agenda in Europe.

Managing and monitoring the ERA partnership

In order to build a sustainable ERA, it is suggested to couple the current top-down policy approach with a bottom–up initiatives in order to get valuable policy feedback and to better identify innovation opportunities and stakeholders needs. Stronger involvement of stakeholders including e-Infrastructures in ERA will improve the decision-making process and policy outputs on the ERA governance. In order to achieve this, establishing an ERA stakeholders' platform and sharing more information with stakeholders via an ERA portal forum is necessary. In addition, establishing dedicated working groups that will discuss specific policy issues will enable more active participation of stakeholders and accelerate the formation of the ERA.

Recommendation 14: The EC should enable stakeholders' participation in the ERA governance and policy development. In addition, the EC should establish an ERA stakeholders' platform and dedicated working groups that will discuss specific policy issues. This should result in the development of an ERA roadmap.

The EC should ensure that the close cooperation seen recently between the Directorate General (DG) for Information Society and Media (the funding source for e-Infrastructure projects) and the DG for Research and Infrastructure (the funding source for ESFRI projects) continues. Better coordination between e-Infrastructures and ESFRI projects will result in increased mutual benefits. Gaining momentum and a strong element of continuity in collaboration between for example EGI

---

[10] http://www.gridforum.org/
[11] http://www.sienainitiative.eu/

and ESFRI projects will lead to stronger links between e-Infrastructures and their strategic user communities.

Recommendation 15: Ensure that there is a closer collaboration between the Directorate General (DG) for Information Society and Media and the DG for Research.

Conclusions

In order for Europe to fully realise its research and innovation potential, the ERA needs to fully accommodate e-Infrastructures into the ERA Framework, or in the words of the EC: "e-Infrastructures foster the emergence of e-Science, i.e. new working methods based on the shared use of ICT tools and resources across different disciplines and technology domains. Furthermore, e-Infrastructures enable the circulation of knowledge in Europe online and therefore constitute an essential building block for the European Research Area (ERA)".[12] Up to now, discussions about the ERA have not paid sufficient attention to e-Infrastructures. However, only when e-Infrastructures are cornerstones of world-class research, can Europe achieve the vision of ERA 2020, i.e., the full accomplishment of "fifth freedom" and "free circulation of researchers, knowledge and technology". Fostering remote access to research facilities, e.g. through the efficient development of appropriate e-Infrastructures will enable faster circulation of knowledge and technology across Europe. EGI is committed to this vision and it needs stronger support from the EC in order to act as a major stakeholder that will give its substantial contribution in enabling it.

E-Infrastructures are now an essential foundation for research and innovation and it is vital that Europe continues to invest in this area at a National and European level. Through its network of federated data centres, EGI aims to provide its user communities with the means to deploy the software environments they need flexibly – where they need them and when they need them. Virtualisation technologies will provide the foundation for this operational model, enabling the research community to access a cloud environment that is tuned to their ever-changing requirements for data-intensive analysis. EC support can provide the means to accelerate the move towards an e-Infrastructure that is able to flexibly and responsively meet the needs of diverse user communities by addressing the following issues described in detail earlier.

Accelerating the development of e-Infrastructures in Europe will be a key enabler of the Europe 2020 Strategy. E-Infrastructures can provide an open platform for innovation where new technologies are created, new services are delivered, new business models are explored and new research results are generated. These innovations will help European researchers harness the data deluge and derive the knowledge they need to provide answers to the societal challenges facing Europe both now and in the future, while contributing to the key actions of the Innovation Union, Digital Agenda for Europe and other key flagship initiatives.

---

[12] http://cordis.europa.eu/fp7/ict/e-infrastructure/home_en.html

Table 1

	Key recommendations for ERA framework
1	EC should provide information (e.g. through Eurostat) on how many researchers need access to distributed digital resources as part of their daily professional activities and how many researchers could benefit from improved access to digital resources.
2	EC should support the establishment of social networking tools to help researchers to connect with each other and advertise/discover research groups. This could be bound to services like OpenAIRE where publications, scientific data and research groups are discoverable and such collaborations used to govern access to e-Infrastructure services.
3	The EC should support applied science communities with funding to develop domain specific platforms to be deployed on e-Infrastructures, thereby offering new innovative services to a community's researchers in other disciplines.
4	EC should disseminate the EGI success story to other national and European RIs in order to demonstrate a possible way of achieving successful cross-border collaborative operation of e-resources for researchers.
5	EC should support the integration of European e-Infrastructures across different resource types (grids of HTC, HPC, petascale, virtualised, data and networking resources) in order to provide a coherent technical and policy structure. If achieved, this integration would enable new capabilities for new innovation within Europe (e.g. multiscale modelling)
6	Establish a specific dedicated technical working group under the ERA framework to help remove the potential barriers and bottlenecks that hinder collaboration between RIs and e-Infrastructures. To disseminate working group's yearly activities, the EC should facilitate an annual forum between RIs and e-Infrastructures that would review the activities of the working group and discuss and adopt its proposals
7	One of the expected outputs from ESFRI projects in the construction and implementation phases should be to formally evaluate the services, capacity and data needed from e-Infrastructures for their research needs and to report this information to the e-Infrastructures.
8	The EC should build an ERA Compendium/Knowledge base with all European RI, ESFRI projects, national and European research groups. A compendium/knowledge base should also describe the assets (e.g. people, services) in the RIs that are part of ERA. Having this structured collection of information would help the research communities to build collaboration and e-Infrastructures and other stakeholders to support their needs. The EC should provide a tool that would enable RIs to update the information anytime.
9	The EC, together with e-Infrastructures, should develop the long-term roadmap for integrated e-Infrastructures aimed at a sustainable operation and evolution of e-Infrastructure services.
10	The EC should tackle the legislative uncertainty around commercial use and collaboration with e-Infrastructures
11	The EC should support EGI in leading the development of a European cloud built from federated virtualised resources from public and commercial providers and not to duplicate and fragment national efforts.
12	The EC should strengthen the relationships between the funding agencies of key foreign countries to develop common policies for better integration of the respective e-Infrastructures. This would have the beneficial side effect of enabling the respective research groups to discover each other and collaborate. Furthermore, researchers would have access to more scientific data and computing capacity as well as to knowledge and expertise
13	The EC should strengthen the relationships with the funding agencies of key foreign countries to develop policies to stimulate the definition of joint standardisation roadmaps encompassing the full standards development life-cycle. This would allow the e-Infrastructure community to remove the remaining interoperability barriers and to provide a seamless set of interoperable services that can underpin the Digital Agenda in Europe.
14	The EC should enable stakeholders' participation in the ERA governance and policy development. In addition, the EC should establish an ERA stakeholders' platform and dedicated working groups that will discuss specific policy issues. This should result in the development of an ERA roadmap.
15	Ensure that there is a closer collaboration between the Directorate General (DG) for Information Society and Media and the DG for Research.