

Below are the comments from the Norwegian grid initiative NorGrid to the EGI Blueprint document v1.2 as distributed to the EGI Policy Board on June 25. NorGrid values the work that has been done to produce the document. Yet, we consider the document to be a draft only that requires further discussion and clarifications. The aim of the comments below is to contribute to improving the Blueprint and make it more acceptable and understandable to NorGrid and presumably similar smaller grid initiatives in other countries.

The Norwegian participation cannot be taken for granted on the basis of Blueprint v1.2. The added value of EGI, the estimates for required man power and the contribution/service fees are most likely considered too vague for the Research Council of Norway to make a decision on providing structural funding available for EGI and how to weigh the direct participation into EGI against its participation in the Nordic Data Grid Facility (NDGF).

1. THE FURTHER PROCESS TOWARDS THE FINAL BLUEPRINT

1.1. The next version of the Blueprint should be available at least 3 weeks before the meeting of the Policy Board in September. Three weeks must be considered the absolute minimum to give all the NGIs and Policy Board members the opportunity to consult with stakeholders and groups. The Blueprint constitutes an important document for EGI (EGI.org + NGIs) and deserves sufficient and broad discussion.

1.2. The next version of the Blueprint (or relevant parts of it) should also be given to the EC for comment. It is important to know EC's views concerning the plans and structure for EGI.org, before a call for tender for EGI.org is issued. This is especially important since the framework for European Research Infrastructures that EC is designing, differs from the EGI.org structure. Also feedback from the EC on the funding scheme will be very useful.

1.3. The next version of the Blueprint should also be given to other EC-funded e-Infrastructure initiatives for comment. These include for example DEISA, PRACE, BalticGrid, SEE-grid, etc, but also regional grid infrastructures like the Nordic Data Grid Facility (NDGF). DEISA is for example using its own middleware and accounting tools and PRACE will have to make decisions concerning these in the Preparatory Phase. It is important that these initiatives actively participate in the design of EGI.

1.4. The next version of the Blueprint must not be the final version. The comments from the Policy Board members (and EC representative) at the meeting July 1 were substantial and addressed fundamental components of EGI. This makes it impossible to have the next version be the final version. At least one more iteration is needed after the next release of the Blueprint.

1.5. EGI_DS should give feedback to each of the NGIs on the comments provided. This to ensure that the comments of all NGIs are indeed given the right consideration and to make the NGIs feel active contributors to the EGI process. In particular, an NGI must be given clarity on issues where the views of the NGI and EGI_DS diverge. It will be harder for NGIs to accept or defend a Blueprint if comments are not taken into account and no further explanation for this is given.

2. COMMENTS ON BLUEPRINT v1.2

2.1. EGI should maintain a list of basic values and principals for all aspects of a pan-European grid infrastructure.

The Blueprint and discussions in EGI_DS could benefit from having a set of principal values for a European grid infrastructure. From these values, the EGI functions can be derived and weighed/prioritized. Some principles are defined for operations and security (Section 5.1.1) but principles are absent for other major functions proposed for EGI.

The basic values must obviously include values that are considered elementary to a grid, e.g., (i) centralized control/activity must be minimized and (ii) sites, organizations, NGIs and user communities must be able to take autonomous decisions and (iii) EGI.org must be independent and unbiased towards technology. Once this list of values exists, it will be easier to discuss the functions and management of the EGI grid infrastructure and in particular weigh the relevance of an activity, decide whether EGI.org or the NGIs should deal with this, etc.

As an example, it is not clear why the central services listed in Section 5.1.4 are mandatory services for EGI.org. Some appear to contradict one or more principals of Section 5.1.1. What is e.g., the added value that EGI.org maintains a central accounting repository, instead of the NGIs keeping such repositories themselves (they have to do it anyway) and report to EGI.org when needed?

2.2. The Blueprint should include some concrete use cases.

These use cases should include examples of how NGIs are expected to interact between themselves, how an NGI interacts with EGI.org (management, operations, fees, reporting, etc.), plus a few real examples of how existing user communities will interact with EGI (EGI.org and NGIs). Such use cases may also reveal which functions and services are mandatory and which not.

2.3. The Blueprint must focus on defining the framework for the long-term EGI infrastructure.

Subsequently, but separately, the Blueprint should describe how to get there starting from today's situation with a manifold of EC-funded projects (EGEE, BalticGrid, SEE-grid, ...). NGIs can then figure out themselves how to get there from their present situation.

It is not always clear from the descriptions in the Blueprint whether a proposed activity is (only) part of the transition phase or will also continue in the long-term model.

The Blueprint does not detail how long the transition period will be, but it is suggested that it may take up to 5 years to build the NGIs properly, see e.g., Section 8.1.3. By then, FP7 is finished, EC will have defined the priorities for FP8, and EGI may have to adjust its Blueprint before the transition period is completed. It will be cleaner to clearly separate the transition period from the permanent EGI structure and define these as two separate phases with a date for when the first (transition) phase must have been completed. Making such separation will also allow for an evaluation at the end of the transition period of the key functions,

management and financial model and what should or should not be continued in the permanent EGI-model.

In Section 8.1.3, the Blueprint does not properly address EU's long-term commitments to funding, e.g., during and after the transition phase, and what the impact will be on the funding required from NGIs. Section 8.1.3 suggests that NGIs will need up to 5 years to get their funding in place (which "makes EC-subsidies essential to get EGI started"), but it is subsequently expected that the EU funding is continued at the same level (20 MEuro/year) after that.

2.4. The Blueprint suggests that EGI.org will act as a controlling body and funding agency for grid development on behalf of the EU.

Two examples:

- Section 5.2.3 states that "the EGI.org Development Unit should be the unique place in Europe where the development of European middleware will be planned and coordinated". At the same time, the Blueprint considers only gLite, ARC and UNICORE to be relevant middlewares. This is too restrictive. EGI must be able to recognize new and innovative middleware concepts.

- Section 8.1.2 proposes that the total membership fee from the NGIs is 1 MEuro to cover the costs for the central administration. The total costs for EGI.org are estimated at 4.5 MEuro (for 45 FTEs). The total requested funding from the EU is nevertheless 20 MEuro. How the difference of 15.5 MEuro is spent is not specified, but presumably considerable part will be used for development of the infrastructure and middleware. Section 8.1.3 includes some support for the proposed 20 MEuro, but it still remains to be seen whether EGI eventually will be as important for science and Europe as GEANT/TERENA, it also remains to be seen whether one general-purpose European-wide grid infrastructure is actually needed.

2.5. The business model for EGI must be improved.

It remains to be seen how many user communities truly will need access to an *entire* pan-European grid infrastructure as envisaged by EGI. There are no numbers in the Blueprint that address this.

The assumed cost of running the total European grid infrastructure is indicated at 76.5 – 84.5 MEuro. It is not mentioned in the Blueprint that this is money well spent. E.g., a valid dilemma for researchers remains the choice between supporting (i) spending 76-84 MEuro per year on infrastructure to couple ca. 40 NGIs and hundreds of existing compute and storage systems or (ii) actually buy a small number of large new compute and storage resources (e.g., for 50 MEuro) and provide adequate operational and user support for these resources (e.g., for 30 MEuro). Alternative (ii) will certainly have much less operational overheads.

Business models that are commercially viable (e.g., models similar to Amazon and Google) are not addressed in the Blueprint.

2.6. EGI_DS must make sure that the EC and all participating countries can and are willing to support the proposed legal organization of EGI.org.

The last paragraph in Section 7.1.2 refers to the legal framework for ERIs that is being defined by the EC and lists a number of advantages for using this framework. The timing by which the ERI regulations are in force and the compatibility with some (unspecified) EGI.org needs are mentioned as reasons to not use of the ERI framework. However, the ERI framework will be the foundation for multiple European infrastructures and thus will be well-constructed and well-understood by all EU-member states (funding agencies) and the EC. The EGI.org organization will be unique and EGI.org and the NGIs will need considerable competence in understanding the legal aspects of this structure (e.g., not mentioned in the Blueprint). Funding bodies may be more willing to support ERI-based structures as they are likely to have higher political status and credibility.

Timing is not a valid argument to choose a legal structure for EGI.org that is different from the ERI framework (or preliminary versions of it). In addition, page 13, section 2.4.1 says that EGI.org should be fast to create. How important is this, considering the fact that EGI.org is supposed to be a permanent entity and that it may take several years for NGIs to satisfy the requirements set by EGI?

The motivation for the proposed legal organization for EGI.org should be detailed in the Blueprint as this is the document on which national funding bodies may base their decisions. One should not expect the funding bodies to read multiple EGI_DS deliverables to understand some of the reasoning in EGI_DS.

Note that PRACE aims to use the ERI framework. EGI_DS must consider whether the use of a different legal organization for EGI may pose problems in the long-term collaboration with PRACE and similar infrastructures (e.g., VAT-regulations, etc.)

2.7. Middleware development should be kept separate from developing and operating infrastructure. Middleware development should not be dealt with by EGI.

Many NGIs have already commented in this earlier. The Blueprint must clearly separate development required for operational tools (configuration database, trouble-ticketing, operational portal etc.) from middleware.

EGI must be open for innovation coming from external sources - also from industry and from outside Europe - and not only promote (internal) development of selected existing middlewares whose quality is subject to an endless debate. Funding of middleware development must happen through open calls and peer-review. It can be doubted that EGI.org and its council is the right body to make decisions on this. European middleware development must not only be EGI/NGI-driven, but also be user/community-driven.

EGI should be *independent* and *unbiased* towards any middleware and software as long as it fulfils certain quality criteria. External development and innovation must always be encouraged. EGI should restrict itself to maintaining a catalogue of essential (mandatory) and desirable services for EGI and promote a set of standards that middleware must or should adhere to. Perhaps the fact that the Blueprint considers only three middleware stacks (gLite, ARC, UNICORE) as relevant is probably sufficient reason to exclude middleware development from EGI.

EGI should not do middleware development unless software for mandatory functionality does not exist and its creation is considered urgent and essential to the functioning of EGI. Selection and validation of middleware components may be considered an EGI-task.

EGI may certify software components to designate that it fulfils specifications and criteria set by EGI (Section 2.3.3), but each NGI should be able to make autonomous decisions on which middleware(s) its operations will be based. NGIs will have to provide services to national or non-European user groups that may have clear preferences concerning middleware.

EGI needs to position grid infrastructure with respect to cloud computing and cloud storage as e.g., provided by Amazon, Google and IBM, and possibly adopt such technologies when this is considered beneficial for the business model of EGI.

The Universal Middleware Toolkit (UMT, page 32) is not well-defined. It reads as if the goal is to merge C++, Fortran, and Perl into one new programming language, but then in a middleware context. Promoting interoperability should be the keyword for EGI. Let each middleware evolve according to internationally recognized standards and target applications areas.

2.8. EGI should interact closely with other initiatives that constitute the European ecosystem.

These include e.g., DEISA and PRACE. The European ecosystem must enable the scaling of applications to use increasingly large resources. The Blueprint seems only concerned with applications that can be split in many independent subtasks. A full-blown European ecosystem must eventually also include scenarios where applications that initially are executed on local compute clusters (hundreds of cores), are promoted to national systems (thousands of cores) and eventually to large European systems as envisaged by PRACE (tens of thousands of cores). Such migration of applications between different layers of the ecosystem requires middleware to get for example the data sets from the users to these systems. The Blueprint is currently not concerned with this. EGI should be independent and unbiased towards any other e-Infrastructure initiative and see what the grid-like needs are that exist in these other e-Infrastructures.

2.9. EGI should promote dynamic resource provisioning and pan-European optimization of the use of compute and storage resources.

One of the real benefits of a grid infrastructure is that peak demands in resources in one country or in one scientific project can be accommodated by temporarily using unused resources elsewhere in Europe. Such scheme for dynamic work load balance of European grid resources is currently not mentioned in the business model of the EGI. The text in Section 8.1.1 mentions that the EGI model foresees that each project partner approaches the NGI of its country and negotiates the availability of resources to the project. This does not cover the possibility to use 'idle' cycles or storage abroad whenever there is a need for this, including 'idle' resources in countries in which this project is not active.

2.10. Management of the EGI (Section 6)

Page 48, Figure. Development (CTO) and operations (COO) are anchored in the EGI management, whereas application support is not. How will EGI.org coordinate application support and how is quality ensured?

Page 51, bullet d. EGI.org and its Council must work with *multi-year* budgets for long-term planning and not only adopt annual budgets.

There is no specification in the Blueprint about how the quality of the infrastructure and its services will be measured. Perhaps this is left to the EGI.org administration and council after they have been established. However, it should be one of the major tasks of EGI.org to measure quality of the delivered services and this should be addressed in the Blueprint.

2.11. Resources and funding (Section 8)

NGIs must be able to calculate (or at least estimate) the contribution fees and the fees of selected services.