

Supporting EGI strategic dissemination activities: The European e-Infrastructures Observatory

Jorge-A. Sanchez-P., Nikos Vogiatzis, Nektaria Berikou¹

JNPartners Co

Xanthou 17, GR15238 Xalandri, Attica, Greece

E-mail: n.berikou@jnp.gr, j.sanchez@jnp.gr, n.vogiatzis@jnp.gr

The ICT-based infrastructures are a crucial asset underpinning European research and innovation policies. They make a major contribution to the objectives of the EU 2020 Strategy and the vision for the ERA, and have a key role in supporting the deployment of new research facilities, whose development is articulated with ESFRI and e-IRG policy groups in a dialogue with Member States.

Designing and implementing e-Infrastructure policies and strategies requires proper knowledge of the status of e-Infrastructures' development in a region/country, including its evolution over time as well as disseminating the results and success stories of respective organizations and individuals.

One of the key challenges faced by e-Infrastructure providers such as EGI.eu, which represents the European Grid Infrastructure of 350 resource centres organised through over 35 national resource infrastructure providers, is to report the diversity and scope of their resources and their usage to funders and decision makers at national and European level.

This paper presents the capacity and functionality of the [European e-Infrastructures Observatory](#)², a single-entry-point and one-stop-shop data warehouse, a yardstick tool that provides to the e-Infrastructures community a comprehensive monitoring and dissemination framework, allowing to keep track of the developments and prospects in e-Infrastructures, including especially a frame of reference for assessing over time the evolution of European Grid Infrastructure(s), and disseminating respective success stories of European-led initiatives.

*EGI Community Forum 2012 / EMI Second Technical Conference,
Munich, Germany
26-30 March, 2012*

¹ Speaker

² The European eInfrastructures Observatory is the objective of the e.nventory project co-funded by the European Commission within the 7th Framework Programme - Research Infrastructures - Grant Agreement no RI-261554. www.eInfrastructuresobservatory.eu

1. Background / Need

Europe has built a world-class operational communication and computing infrastructure to facilitate scientific research. These e-Infrastructures provide scientists throughout Europe with very high-speed communication networks and powerful computing resources that foster collaboration and enable front-end research.

EU support to research infrastructures as a whole and to e-Infrastructures in particular, has been -and continues to be- substantial by any (international) standard and by any comparison, setting European research at the forefront of the policy agenda and reaffirming the priority that Member States place on research as the foundation of European competitiveness.

To facilitate a rapid transition to e-Science, the European Commission and Member States have made significant investments in e-Infrastructures, including e-Science grids, the pan-European research network GÉANT, data infrastructures and supercomputing. These significant investments have been materialised through primarily the so called “flagship projects” such as [EGEE/EGEE-II/EGEE-III/EGI-InSPIRE](#) (access to distributed computing platforms / e-Science grids), GN/GN2/GN3 (high speed research networks - [GÉANT](#)), [DEISA](#), [PRACE/PRACE-II](#) (developing supercomputing capacity for very demanding applications), and new initiatives to support the emergence of a robust platform for access and preservation of scientific information (Scientific Data eInfrastructure - SDI), Global Virtual Research Communities (GVRC) and so on [6].

There are also many other valuable related e-Infrastructure projects that address new and different user communities and help to develop and diffuse this technology to ensure its exploitation is optimized, like [SEEGRID](#) that helped extend grid infrastructures to regional communities in SE Europe as well as [EUMEDCONNECT](#) or [SEEREN](#) that helped similar geographical extensions of R&E networks.

The aforementioned projects and initiatives often embedded in their work programme several sub-activities dealing with evaluation. However, these evaluations were primarily focused in a sub-domain of e-Infrastructures as per each respective project as well as being often contained geographically within the project partnership.

Yet, designing and implementing e-Infrastructure policies and strategies requires thorough knowledge of the status of e-Infrastructures’ development in a region/country, including its evolution over time as well as its impact due to its use by organizations and individuals. Consequently, calls for e-Infrastructure development statistics have become more frequent, both at the national and the international level.

As a matter of fact, the demand for e-Infrastructure statistics has increased significantly as public organisations seek to design, monitor and review national policies and strategies to take advantage of the rapid advances in information technology, while the private sector (business communities) exhibits interest and requires information about access to, and use of, e-Infrastructures by various consumer groups, as well as information about the impact of e-Infrastructures. At the international level, comparable e-Infrastructure indicators are critical to cross-country comparisons of e-Infrastructure development and for monitoring the global digital opportunity (or divide) in this respect.

The issue of benchmarking has also been raised high in the European policy agenda in several recent occasions as vital for assessing the impact of eInfrastructures at EU level and beyond [1], [2], while the FP7 Research Infrastructures Work Programme 2010 materialized policy into specific call-for-action, addressing “specific studies on e-Infrastructure related topics, in particular to evaluate the impact of the e-Infrastructure programme including the establishment of appropriate indicators”[3].

As highlighted in the Communication "Responding to Strategic Needs: Reinforcing the use of evaluation", a rigorous evidence-base is crucial in decision-making for planning, designing and implementing EU policies [4].

During the 8th e-Infrastructure Concertation Meeting (2010), the key players in the e-Infrastructures community, working towards a long term sustainable e-Infrastructure for scientific research in Europe agreed to work towards the development of a world-class resource as part of a strategy to achieve European leadership in e-Science as well as to take stock of progress in e-Infrastructures and to discuss impact assessment in the domain [5].

Furthermore, in the most recent impact assessment study of e-Infrastructures in FP7 it was revealed that:

- A periodic, project-based assessment that can be aggregated at overall Programme level is necessary
- A holistic approach and supporting tools are missing, able to
 - monitor eInfrastructures impacts and trends at EU/MS (and beyond) level and reveal interrelations to macroscopic socio-economic indicators;
 - utilise “factual”, unbiased metrics and data without subjective interpretations;
 - apply a validation/curation process to ensure that the utilised data accurately represent the eInfrastructures status quo;
 - take into account owners, users, and sponsors of eInfrastructures at both EU and National level;
 - cover regions beyond EU with the same monitoring framework thus allowing for cross comparisons of EU eInfrastructures to international ones.

The European Commission has declared a clear need for assessing the EC investments in the area of e-Infrastructures, a need for reaching a common understanding, stimulating and engaging key stakeholders, raising awareness and getting feedback on the “how” and “what” of upcoming socio-economic evaluation and specifically of impact assessment of e-Infrastructures.

Following the investment in e-Infrastructures and responding to the e-Infrastructures community’s strategic need for a tool to aid evaluation, both prospective and retrospective, of related e-Infrastructure initiatives and programmes, the [European e-Infrastructures Observatory](#) [8], as elaborated in the following sections, is able, through the collection and utilisation of appropriate indicators, to monitor the development and prospects in European Grid as well as other key e-Infrastructures, and through an intuitive, interactive and user-friendly interface, to communicate all findings to related stakeholders and to the public-at-large, in a seamless and impartial way.

2. Methodology

The work towards establishing the European e-Infrastructures Observatory was structured so as to achieve: (a) identification of a core set of benchmarking indicators for the European e-Infrastructures Observatory that will be the baseline for monitoring e-Infrastructures development progress, i.e., collecting, measuring and interpreting indicators related to e-Infrastructures, (b) e-Infrastructures stakeholders' feedback and consensus on the proposed structure and functionality of the European e-Infrastructures Observatory, i.e., verifying, validating and updating the data values for such indicators from various sources, and (c) European e-Infrastructures Observatory functionality demonstration through a prototype web platform that will be available to e-Infrastructure communities and to the general public.

The baseline list of indicators is a core element of the European e-Infrastructures Observatory monitoring framework, thus their selection in conjunction with the consultation activities is a fundamental process for the success of the work.

Developing 'good' indicators is not an easy task; it involves collection, collation and systematization of data. Their development and selection often result in long lists of indicators selected on the basis of subjective perception. Such lists tend to treat some topics in depth while others are ignored. Having too many indicators can also result in confused priorities and overwhelming details for both developers and users. Many of these problems can be avoided by using a diligent development method, comprised of discrete and clear phases.

For the case and purpose of the establishment of the European e-Infrastructures Observatory, the development process was broken down into five distinctive phases in order to plan its implementation in a more effective way.

Phase 1: Establishing the purpose of the indicators

A critical step in defining the suite of indicators was to clearly identify the target audience and purpose for the indicators. This determined the scope of the indicator set and assisted in keeping the overall activity focused and in perspective. Key questions / decision points that determined this process were:

- Who will be responsible for the final selection and publication of the indicators?
- How will key stakeholders be involved?
- Will a group of experts be established to provide specialist advice?
- Will public consultation be undertaken?
- How will the indicators be sustained and funded over time?

The purpose of measurement and the process of deciding what to measure and determining who will benefit from the indicators, were also as important as the issues of what to measure and how to define specific indicators and technical methods.

Phase 2: Grasping the conceptual framework

A "conceptual framework" to guide the development of the set of indicators was conceived. The absence of a framework can result in the generation of an eclectic mix of indicators, with no clear rationale for their selection. The framework provided a formal way of

thinking about the topic area, ensured the selection of indicators is relevant and balanced and it helped the understanding of the complicated links between the indicators, and provided a useful device for organising and reporting on indicators in a structured and meaningful way.

In this respect, a thorough study was carried out to examine international best practices for similar research and innovation observatories in a wide variety of fields and domains. A cartography of online and offline e-Infrastructure directly or indirectly related databases and repositories around the world was also delivered, to identify related databases where raw data values are stored and can supply the indicators that will monitor e-Infrastructure development or that can be useful to exhibit correlations with e-Infrastructure-related indicators.

Phase 3: Selecting the indicators

The selection process of indicators took into account the purpose of use of the indicators, and the expectations of the information sought to be obtained by each target group involved in the process. The indicators meet a whole set of criteria, which both condition and limit the way they are developed, constructed and used. The principles that drove the selection of the indicators were:

- Measurable and quantifiable: reflect the phenomenon to be measured and provide guidance that is reliable and scientifically defensible and potentially valid to compare depending on the purpose of the indicator.
- Meaningful: appropriate to the needs of the user, so as users agree with and accept the indicators for gauging presence of and change in values.
- Clear in value and in content: distinct indication on which direction is good and which is bad as well as measured in understandable units that make sense.
- No redundancy or double counting: ensure that indicators do not overlap in what they measure.
- Feasible: measurable at reasonable cost and effort.
- Comparable: over time and also from one location to another. Usually, if the indicator is reliable or verifiable, it is also comparable.
- Comprehensive: sufficiently describe all essential aspects of the system under study.
- Timeliness: be quick and easy-to-produce since there is no point in having indicators that take long to measure and thus they are no longer relevant. Last, but not least, there should be minimal time lag between the collection and reporting of data, to ensure that indicators are reporting current rather than historical information.

Phase 4: Interpreting and reporting indicators

The interpretation and reporting of indicators is a critical stage, as it bridges the gap between measurement and understanding. In deciding how to report the indicators, the target audience should be kept in mind. A mix of graphs and commentary is generally more effective for a public audience than large amounts of commentary. The presentation of the indicators is kept objective and policy-neutral. Contextual information, such as changes in key demographic variables, is provided where appropriate to assist the interpretation. All aspects of the indicators are clearly transparent.

An in-depth study of available international data observatories as well as of indicators visualisation and representation modelling and techniques took place in order to collect best practices on how to approach and construct a top-class European e-Infrastructures Observatory, including its structure, methodology and data organisation approach.

Phase 5: Maintaining and reviewing the indicators

Indicators should be subject to an on-going assessment. They should be open to challenge, discussion and modification to reflect changing objectives, the emergence of new issues and the improvement in measurement techniques and data availability.

The project has established a formal mechanism for reviewing the indicators. This involved open consultation with stakeholders, including technical and subject-matter experts, data providers, the target audience for the indicators, other interested groups and the community.

The entire consultation process of the project has been designed by taking into account this very last statement. It was important to listen to the feedback provided through the review process and to react by making adjustments to the indicator report. Sufficient liaison took place with key e-Infrastructure stakeholder organisations and individual experts, along the lines of engaging with an active Advisory Board.

Such key e-Infrastructure stakeholders included key European e-Infrastructure flagship projects, such as the Integrated Sustainable Pan-European Infrastructure for Researchers in Europe (EGI-InSPIRE). A Memorandum of Understanding (MoU) has been signed with the European Grid Infrastructure (EGI) with the aim to establish a solid and sustainable framework for monitoring the progress of Grid infrastructures and better respond to the strategic need of the Grid community to properly leverage and disseminate EGI's role and contribution at National and Pan-European level.

3. Results

The objective of the European e-Infrastructures Observatory is to develop a reliable framework that can be used as a yardstick by e-Infrastructures stakeholders (i.e. the European Commission, the European Parliament, national governments and research funding bodies, e-Infrastructure policy bodies and projects, scientific/research communities) for tracking and detecting trends and –potentially- initiating corrective actions as well as supporting the designing of new programmes for future development of e-Infrastructures.

The European e-Infrastructures Observatory has reached a mature functional state that portrays a multifaceted, multidimensional and polymorphic panorama of e-Infrastructures in Europe:

- Features a core set of 45 benchmarking indicators that is the baseline for monitoring e-Infrastructures development progress, usage, impact and investments drawn;
- Provides access to more than 10.000 individual figures and other related data and benchmarks;
- Offers 7 intuitive, interactive and user-friendly visualisation tools based on best practices in visualisation and representation techniques;

- Consolidates more than 18 months of stakeholders' feedback and consensus on the structure and functionality of a world-class e-Infrastructures Observatory.

3.1 The European e-Infrastructures Observatory indicators

The process of developing the core set of indicators followed the framework of five (5) distinctive phases defined in section 2 "Methodology" that includes the process for collecting, measuring and interpreting indicators related to e-Infrastructures as well as the process of verifying, validating and updating the data values for such indicators from various sources in conjunction with the consultation activities. The baseline list of indicators that makes up the core of the European e-Infrastructures Observatory monitoring framework is the final outcome of a thorough selection process that took into consideration several related factors and achieved consensus within e-Infrastructures' stakeholders.

Some fundamental reference points and prerequisites for the selected indicators are the following:

- They are able to capture the core e-Infrastructure building blocks (i.e., *grids, networking, supercomputers*), including their key development aspects such as *capacity, ease-of-access*, and *utilization* of digital resources and technologies, as well as measuring e-Infrastructures penetration and adoption across several dimensions, such as *infrastructure-based elements*, network quality characteristics, and knowledge-intensity of the R&E population, and enable differentiation between them.
- They can be measured and verified with relative ease, being concrete and readily understandable by stakeholders in order to facilitate their decision-making, and can ultimately be maintained and updated via consistent and straightforward methods.
- Though diverse and covering a wide range of areas, the indicators are grouped in consistent classifications/categories/areas where special interest exists for the e-Infrastructures stakeholders.
- They fit the purpose of wide deployment, having in mind at least the EU Member States to start with but are also able to cover an extended number of countries and regions, without the need of complicated adaptations.
- They are based on reliable data and widely-approved metrics in order to have agreement and consensus in the e-Infrastructures community and its stakeholders that the indicators registered are serving their purpose well.

A first set of indicators has been identified, including Networking, Grid and Supercomputing indicators, plus a set of additional supporting indicators for cross-correlations and has been used to consult with stakeholders and communicate to the collectors of the indicators the need to supply the related indicator databases. In this respect, meetings have been held and input has been received from [TERENA](#), [EGI.eu](#) and [PRACE RI](#). As far as the additional indicators is concerned, there was a vast wealth of options available, stemming from the long history of organisations and initiatives on benchmarking and development monitoring, such as the [ITU](#), [WorldBank](#), [UNESCO](#) and [Eurostat](#). Data were freely accessed and reused, providing that acknowledgements are duly made to the owner/source. An exception applies to

the ITU related indicators that was handled with a request for permission to use the data published in the World Telecommunication ICT Indicators database [9].

All indicators and related data/values have been extensively tested against the framework defined and extensive data “curation” was performed in both the e-Infrastructure data received since it was noted that verification of collected data was not systematically performed even by the most mature e-Infrastructure data repositories as well as in the additional indicators that were extracted from the respective databases (ITU, UNESCO, Eurostat, Worldbank).

During this process, it has been proved that the aspects of time-span and geographic coverage are quite challenging to deal with in a single, horizontal way: in several cases the time period for which data is available may appear to be quite extended, however partial data may be available for some (usually earlier) years; likewise, geographic coverage, even in cases where it addresses in principle all regions and countries, several incomplete cases may be included in the whole list.

The availability of data and the overall capacity played also a central role in the quality and suitability of each indicator. Some of the indicators were also considered to be difficult to use, others did not meet the abovementioned requirements or the work they required in the data collection process was not an easy and clear task.

The issues of different levels of “quality” with respect to the available data for the core building blocks and in certain cases time-span and/or geographic gaps make it challenging to obtain complete and consistent data for all observed indicators.

After this revision process all “non-compliant” indicators were removed and another round of contacts took place with the source/owner of the indicator for which missing data has been detected which helped to rectify the situation (though still not at a full coverage).

The final set of indicators comprises of thirty eight “numerical” indicators reorganised in ten groups so as to better reveal their capacity: Networking infrastructure, Supercomputing infrastructure, Grid infrastructure, Usage, Users, Investments, Impact, Digital agenda, Europe 2020, Other/Supporting.

Currently the European e-Infrastructures Observatory covers the timeline 2001-2011 but the collection of new values of the indicators is an ongoing procedure. It is expected that the data will be updated twice a year taking under consideration the availability of new values from the original sources.

Table 1 summarizes the indicators that were selected and adopted for populating the European e-Infrastructures Observatory. The data for all the indicators has been collected from the following data repositories: TERENA, EGI.eu, TOP500.org, ITU, Eurostat, Worldbank and UNESCO.

Table 1. Set of European e-Infrastructures Observatory indicators

A/A	Indicator	A/A	Indicator
1	NREN ³ typical backbone capacity	20	Gross Domestic Expenditure on R&D
2	NREN direct IP connection to GEANT ⁴	21	Patent applications
3	National supercomputers maximum performance	22	Publications
4	National supercomputers peak performance	23	Fixed broadband penetration
5	National supercomputers number of cores	24	Students in tertiary education
6	National supercomputers number of processors	25	Students in secondary education
7	National supercomputers in the top500 list	26	NREN budget
8	NGI ⁵ number of cores	27	Unemployment rate
9	NGI installed disk capacity	28	Mobile broadband penetration
10	NGI maximum performance	29	Regular internet users
11	NGI number of sites connected to the NGI	30	Population who have never used the internet
12	NREN IP traffic from external networks	31	Fixed broadband Internet connection charge
13	NREN backbone congestion index	32	Fixed broadband Internet monthly subscription
14	NGI core hours	33	International commercial Internet bandwidth
15	NGI number of jobs	34	Gross Domestic Product
16	Researchers	35	Gross Domestic Product per Capita
17	R&D personnel	36	Gross enrolment ratio
18	Teaching staff in tertiary education	37	Total Population
19	Teaching staff in secondary education	38	Surface area

The European e-Infrastructures Observatory also exhibits European achievements via an additional set of metrics found in the “Stakeholders mapping”, the “Pan-European Networks” and the “Chronology mapping” services that are further elaborated in the following section.

In order to better coordinate actions with related fellow projects, collaborations with on-going initiatives endorsed by the European Commission a) *performing economic evaluation and monitoring in the area of eInfrastructures*, namely ERINA+⁶ and RI-Impact⁷, b) *mapping the*

³ NREN: National Research and Educational Network

⁴ GEANT: the pan-European data network dedicated to the research and education community

⁵ NGI: National Grid Initiative

⁶ <http://www.erinaplus.eu>

⁷ <http://www.ri-impact.eu/ri-impact/index.php>

European Research infrastructure landscape, namely MERIL⁸, e-Infranet⁹, EuroRis-Net¹⁰, were also explored. An area of mutual interest in the data collection process and in the difficulties encountered for the gathering of some of the data was identified and the projects explore the possibility to identify a set of common recommendations based on the experiences and results obtained by each project.

The information/consultation/validation process with key individuals, projects, initiatives and e-Infrastructure stakeholders will continue for a group of benchmarks and visualisations that are still under discussion and examination and will be added in the near future.

3.2 Services on the European e-Infrastructures Observatory

The European e-Infrastructures Observatory services are accessible via a web portal and organised in two levels that currently allow for:

- access to an advanced set of monitoring, dissemination and collaboration tools, including geographical mappings, graphs (bar/line/bubble charts), scatter plots, hierarchically-structured representations, custom-build composite indices allowing to examine interrelations between indicators; timelines of e-Infrastructures success stories;
- access to a time series trends service, that displays on a time series graphs, data for all countries over time enabling the identification of the countries that have improved or are still at an early stage in terms of Grid and other blocks of e-Infrastructures over the past 10 years
- access to eagle-eye overviews of national grid initiatives (NGIs) currently established and active all over Europe; global e-Infrastructures organisations, interconnections of the grid and global scientific communities;
- access to specific e-Infrastructures intelligence (developments and trends) through resources monitoring and infrastructure status quo;
- easier/quicker identification of e-Infrastructures organisations in Europe and beyond, that possess particular competences of interest;
- increased visibility of the EU-led, EU-supported and of EU-interest e-Infrastructures internationally.

The usage of the European e-Infrastructures Observatory is constantly increased. The most popular services so far are the “Indicators on Maps” (53% of the people that visited the Observatory during the last two months used this service) followed by the “Stakeholders mapping” service (49% of the people that visited the Observatory during the last two months used this service).

A short description and a figure for each visualisation tool/service follows.

3.2.1 Indicators on Maps

It's a service to depict indicators on a geographical map, as well as graphs allowing the user to examine interrelations between indicators. The map is centred on Europe with the range of indicator values presented with colours. The tool allows the user to view a list of all available

⁸ <http://www.esf.org/activities/science-policy/research-infrastructure/meril-mapping-of-the-european-research-infrastructure-landscape.html>

⁹ <http://e-infranet.eu>

¹⁰ <http://www.euroris-net.eu>

indicators, countries and years and be able to select one. A “playback” option is also available to view indicator progress per year in animation. The map appearance can be changed by the user. Export options are also available: the indicator data may be exported in a table or on a map.

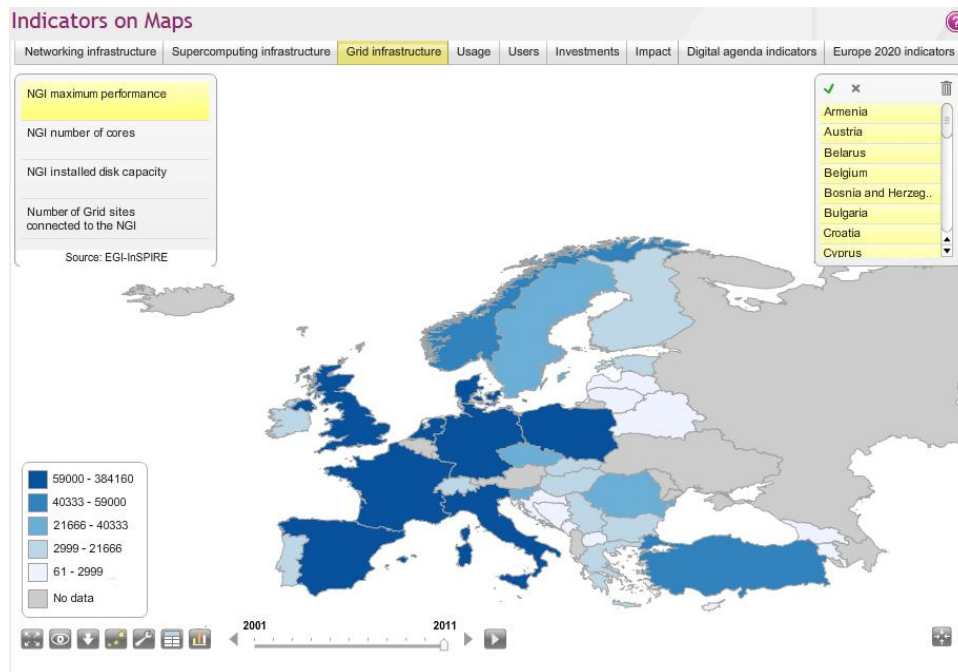


Figure 1: Indicators on Maps service

3.2.2 Trends

This service displays, on a time series graphs, data for all countries over time. The user may select particular indicators and countries and observe trends on how their values are changing over time. The user can identify the countries that have improved or are still at an early stage in terms of e-Infrastructures over the past 10 years. The countries as well as the indicators can be categorised and the user has the option to display all of them or a selection. The Trends service is also appropriate to generate horizontal or vertical bar charts. Animation is possible with the animation button in the time slide. Last but not least, the Trends service is suitable for scatter plots, a very useful graph to identify correlations between different indicators.

A “logarithmic scale” button is also available that may be useful especially for indicators such as data storage or number of CPUs for which the linear view does not offer the best representation.

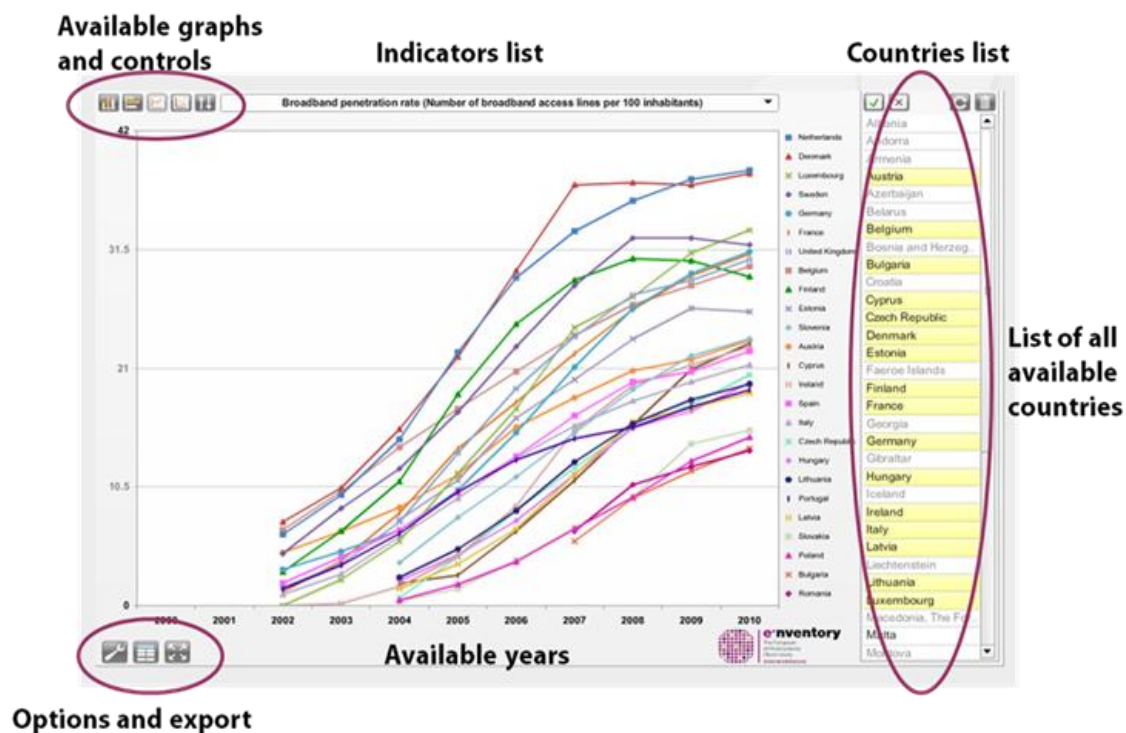


Figure 2: Trends service

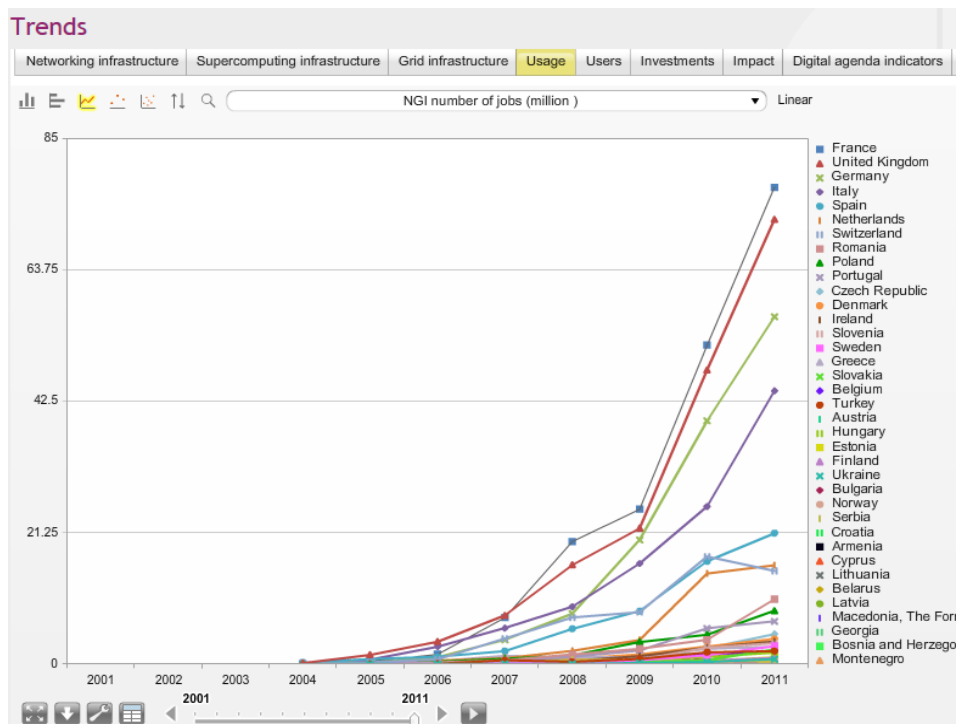


Figure 3: NGI number of jobs (series graph)

3.2.3 Treemaps

This is a non-geographical spatial map service for displaying indicators that are hierarchically-structured as a set of nested rectangles. The service allows the user to view a list with all the available indicators and be able to select one. The following figures display the NGI number of jobs on the treemaps service.

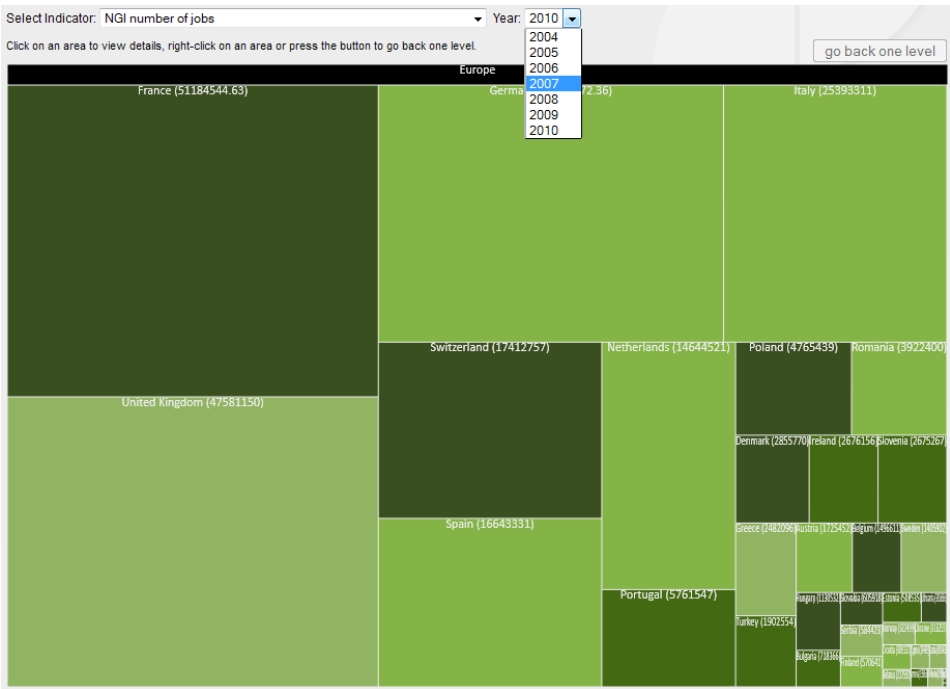


Figure 4: Treemaps service, NGI number of jobs (1st level)

The user may then click on any of the countries to see the details of each site connected to the NGI (2nd level, Figure 5). To go back up one level, the user can click on the respective button above the treemap. Then clicking on any of the sites the user may see the details of the respective indicator (3rd level, Figure 6).

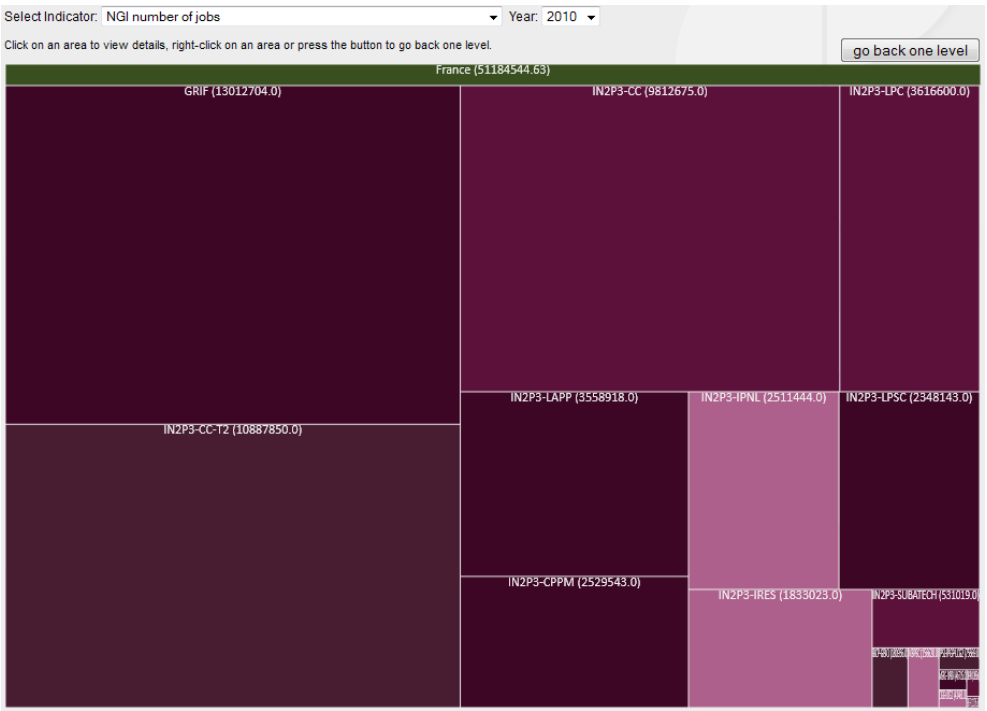


Figure 5: Treemaps service, NGI number of jobs (2nd level)

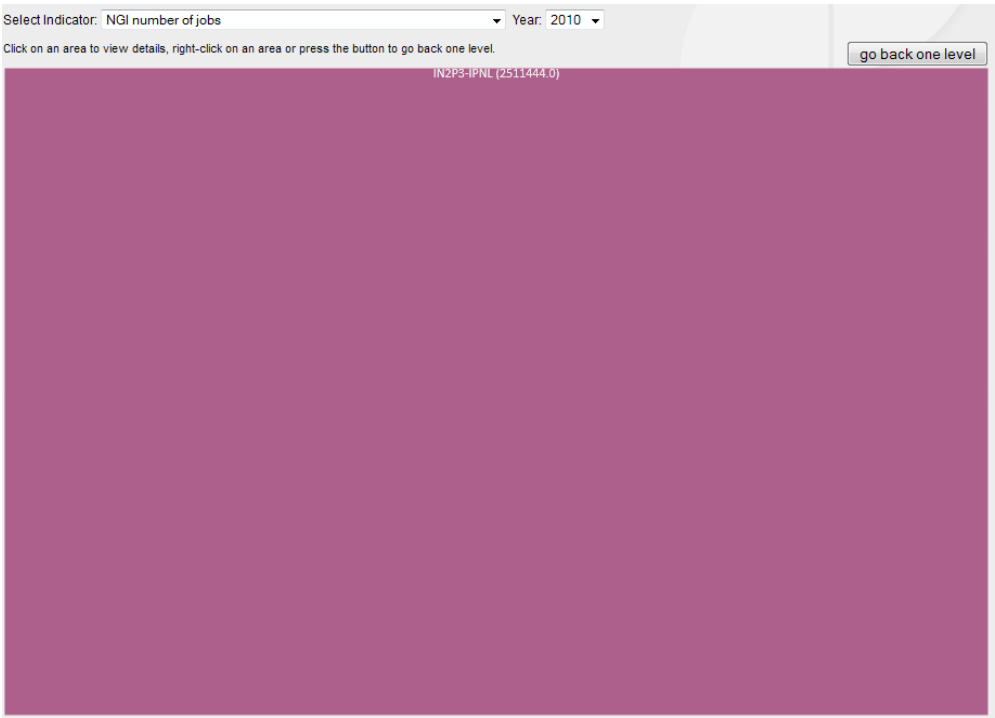


Figure 6: Treemaps service, NGI number of jobs (3rd level)

3.2.4 Stakeholders Mapping

It is an eagle-eye mapping overview of the e-Infrastructure organisations currently established and active all over Europe. The user can choose to display four different types of organisations on a map: National Grid Initiatives, National Research and Educations Networks, Supercomputer Centres, and Coordinating Entities. The user can filter the organisations by their type and also may select one or more countries and/or regions of interest. Detailed information is available for each organisation. The tool supports visual cues that are distinguishable by user with colour vision deficiencies.

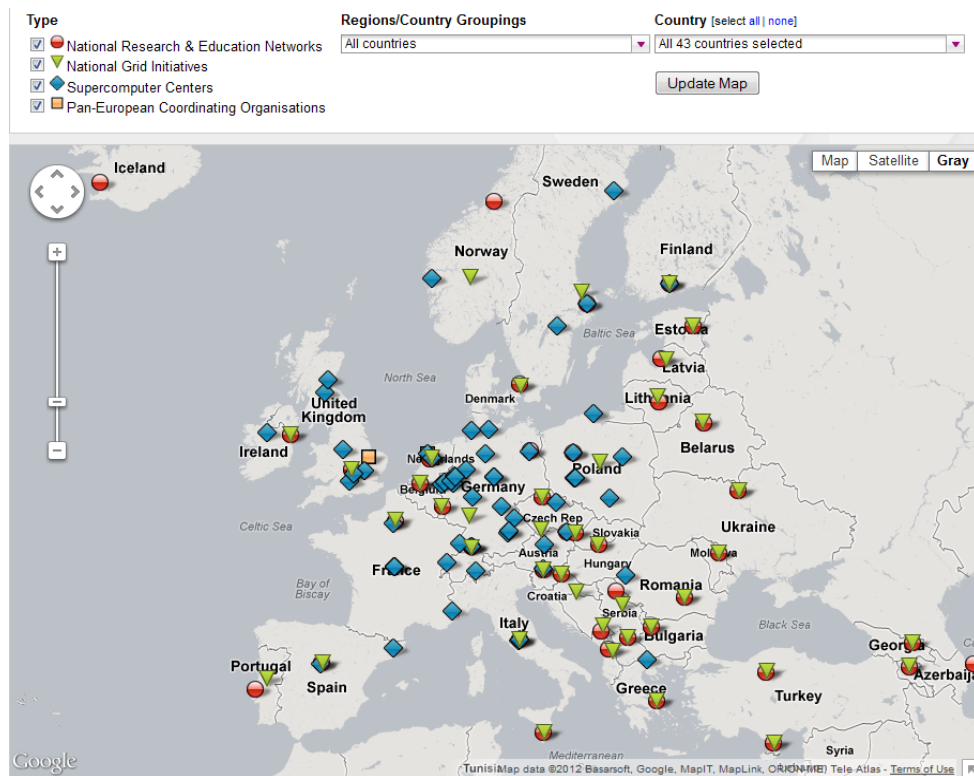


Figure 7: Stakeholders Mapping service

3.2.5 Pan-European Networks

It is a graphical service showing, with coloured lines on a map, the interconnections of scientific communities and networks. The service allows the user to observe the evolution by clicking on the play button and watch the changes from a start point in time until the most recent point in time for which data is available.

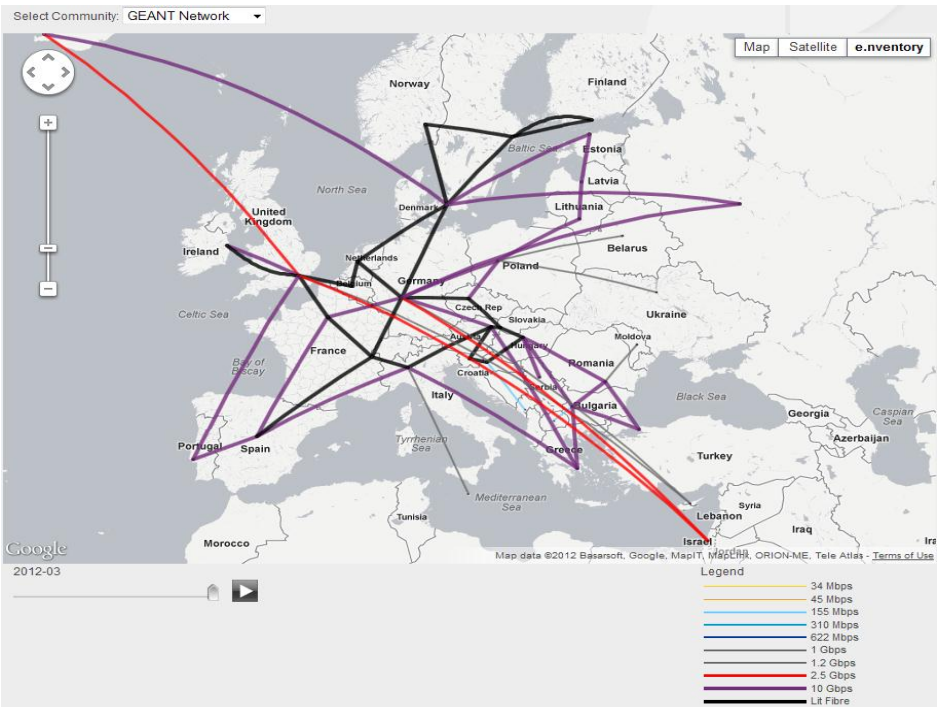


Figure 8: Pan-European Networks service

3.2.6 Chronology

The Chronology service is a timeline of major events in the history of e-Infrastructures. Events are presented on a timeline. The user may focus on particular periods and select events that are of interest, to view their description.

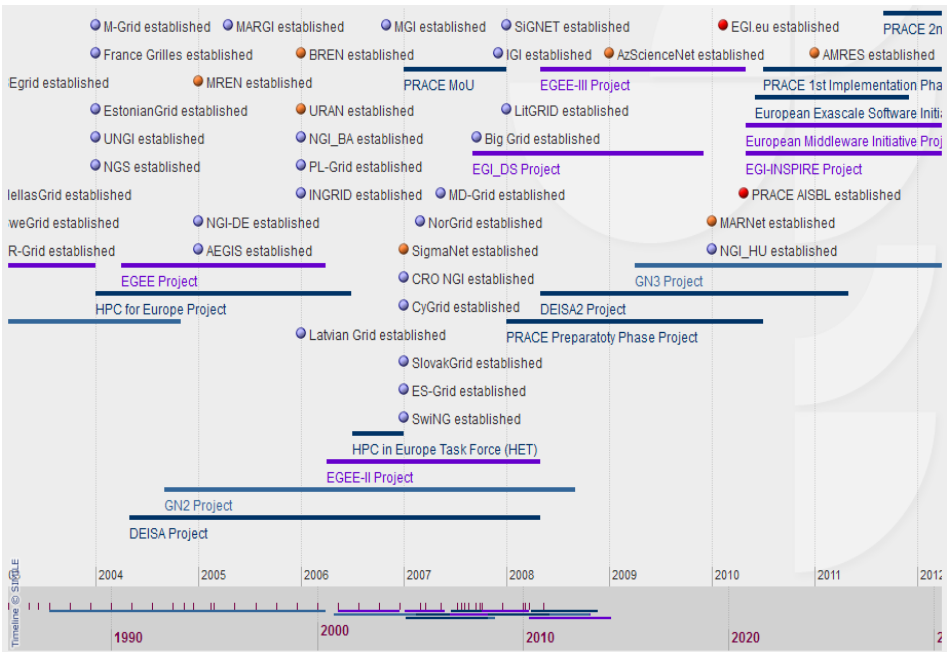


Figure 9: Chronology service

3.2.7 Build Your Own Index

This service allows the user to custom-build composite indices and to create unique country comparisons and rankings by applying, with relative freedom, formulas and weights to the available data. The user may view a list with all the available indicators and be able to select one or more. For each indicator it is possible to set a weight (0-100%). Then specific years may be selected as well as a normalisation algorithm and a name for the composite index. The composite index is stored during the session.

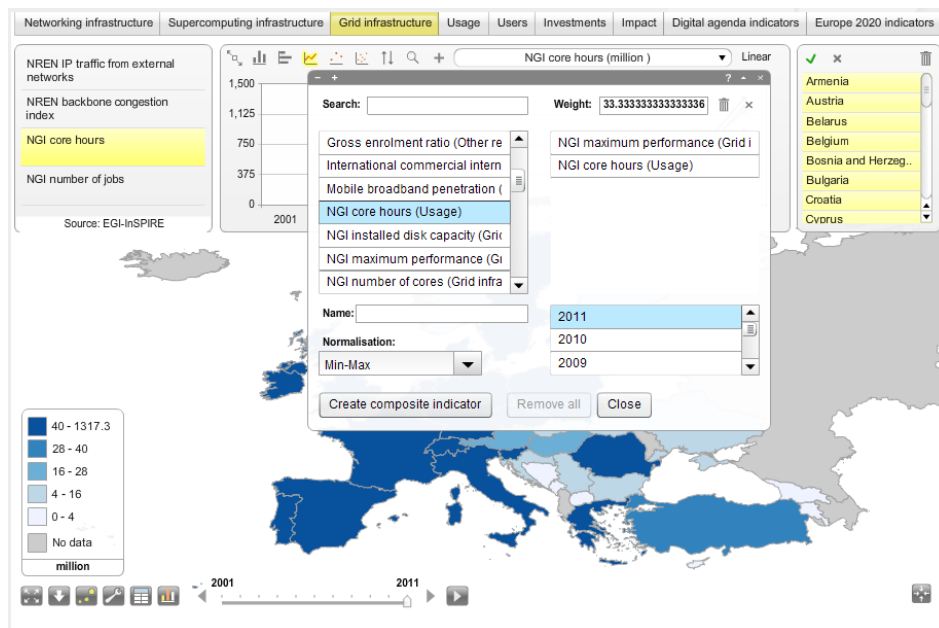


Figure 10: Build Your Own Index service

3.3 Majors problems encountered and lessons learnt

Although several observatories already exist for several disciplines that provide access to a variety of data, none possess the consolidation and unification level that may cover the full spectrum of eInfrastructures. On the Data Repository front, and contrary to the Observatory scene, there are quite a few databases with stored data that are collected regularly.

Among the most mature ones, the TERENA Compendium of National Research and Education Networks in Europe [10] is accepted by the wide NREN community as a reference point for data gathering and archiving. It constitutes the leading survey mechanism for collecting data on NRENs, in a consistent way, over a period of time. NRENs from outside Europe are invited to submit data as well. The survey is conducted in a transparent way and the answers to the survey are made public as soon as they become available.

However, although one-stop-shop data repositories that concern NRENs are available, the same does not apply for grid- and super-computing. There are different levels of “quality” with respect to the available data. More precisely:

- On the networking side, the organisational maturity that has been reached all over Europe for NRENs after a sufficient period of time and a consecutive series of flagship EU projects (currently in its GN3 implementation), coupled with the presence of a dedicated data collection mechanism which is carried out through a single, well-established and acknowledged related organisation (TERENA). The TERENA Compendium is responsible for a substantial amount of available indicators that exhibit a high degree of reliability in terms of time-span and geographical coverage.
- On the grid computing side, things seem to be on a transitional period, with several sources of data available containing a significant range of data and statistics, but sometimes lacking consistency in terms of time-span and geographical coverage (when compared to what the TERENA compendium has achieved). A one-stop-shop convergence can be expected to take place soon through EGI and EGI-InSPIRE (and potentially with the support of the European e-infrastructure Forum). This should lead towards a single, quantitative and qualitative point of reference for Grid progress monitoring that will enable a unified and seamless access to the Grid state-of-play to be made available. Within the framework of the Memorandum of Understanding (MoU) which has been signed between the European e-Infrastructures Observatory and the EGI various NGI Key Performance indicators were defined to track the evolution of grid infrastructures across Europe in terms of resources contributed and funding levels. Substantial feedback was also generated on possible improvements in the data collection mechanism, as well as on new or revised indicators for which the collection of data may be important. The collaboration and exchange of ideas is on-going with the aim to establish a solid and sustainable framework for monitoring the progress of Grid infrastructures.
- On the supercomputing side, the respective data aggregation framework is still at its infancy as far as an EU-hosted initiative is concerned. Besides the top500.org initiative, there has been little progress from the European side in terms of setting up a data repository mechanism. It is anticipated that the following phases of the flagship PRACE project could lead such an initiative and could help develop a single data collecting mechanism for supercomputing infrastructures. A round of contacts with PRACE took place so as to communicate those findings, exchange ideas and discuss on possible improvements in their data collection mechanisms and in their communication with the stakeholders.
- On other emerging e-Infrastructures (e.g. data infrastructures, cloud computing) both data aggregation as well as appropriate indicators, are still at their infancy, if not inexistent.

An overall concern about the validation and collection of data has also been expressed during the consultation activities. Verification of collected data is important but it is currently not systematically performed even by the most mature e-Infrastructure data repositories. In order to achieve a credible and consistent process of data collection and validation, this has to come as a request from authorities that actually finance the eInfrastructure organisations, otherwise there is a lack of leverage.

In addition, the aspects of time-span and geographic coverage are quite challenging to deal with in a single, horizontal way: in several cases the time period for which data is available may appear to be quite extended, however partial data may be available for some (usually earlier) years; likewise, geographic coverage, even in cases where it addresses in principle all regions

and countries, several incomplete cases may be included in the whole list. Combined, time-span and geographic coverage gaps, make it challenging to obtain complete and consistent data for all observed indicators.

To address current gaps and hurdles and evolving the state of affairs towards a strategic, sustainable and holistic framework for keeping track of e-Infrastructures development over time and global regions of interest, the following noteworthy items are currently missing:

- A periodic, project-based assessment that can be aggregated at overall EU Programme level.
- A holistic approach for consistently applying data validation/curation in order to ensure that the collected data accurately represent the e-Infrastructures status quo.
- A set of data points that would enable to cover regions beyond EU with the same monitoring framework thus allowing for cross-comparisons of EU e-Infrastructures to international ones.
- A set of indicators and benchmarks to monitor emerging e-Infrastructures such as data infrastructures, cloud, as well as other areas of interest.
- A universal framework for monitoring e-Infrastructures impacts and trends at EU/MS (and beyond) level and revealing interrelations to macroscopic socio-economic indicators by utilising “factual”, unbiased metrics and data without subjective interpretations.
- A collaboration framework that would be able to accommodate the views of owners, users, and sponsors of e-Infrastructures at both EU and National level.

4.Future work

European Commission’s positioning on the importance of Impact Assessment and the exhibited general audience interest at the 8th and 9th e-Infrastructure Concertation Meetings, the list of e-Infrastructure High-Level Experts participating at the Advisory Board, and the 19 Letters of Support that were received in November 2011 from global e-Infrastructures stakeholders with regard to the continuation and the expansion of the project, inspire confidence that the e-Infrastructure community do find this activity to be of significant added-value.

The results achieved to this date have generated awareness among e-Infrastructures stakeholders with regard to using as well as developing further a (European) e-Infrastructures Observatory. Apart from European and National institutions that are sponsoring electronic infrastructure initiatives and Policy Bodies that have a clear mandate of supporting the development and the sustainability of electronic infrastructures, stakeholders that have declared their interest in e-Infrastructures Observatory developments include flagship EU initiatives such as GEANT, PRACE, EGI, [EUDAT](#), [OpenAIRE](#), as well as global e-science frontrunners from the USA, Russia, China, Australia, South Africa, Taiwan and Latin America.

Recognising the importance of sustainability and the steps that can lead to efficient utilisation and exploitation of the European e-Infrastructures Observatory, a second phase is envisioned that aspires to include:

- Maintenance of the services developed and constant update of the data of the indicators that makes up the core of the European e-Infrastructures Observatory monitoring framework;
- Studying additional types of e-Infrastructures, such as Data e-Infrastructures /repositories and Cloud Computing. In this regard, an awareness campaign will be initiated to introduce the

platform to these specific areas by addressing the groups handling them, capitalising in particular the planned collaboration with the e-Infranet project which aims to build a network that will develop and strengthen cooperation and coordination between national e-Infrastructures in the fields of Cloud computing, “green” computing, data Infrastructures (incl. open access) and other new/emerging eInfrastructures;

- A follow-up EC-sponsored project that will pursue the geographic and thematic expansion of the observatory towards global coverage of established and emerging e-Infrastructures;
- Cross-licensing of the developed platform technologies and tools to other monitoring and benchmarking initiatives and projects, either within the e-Infrastructures field or in completely other domains. The way of collaboration with e-Infrastructures projects and initiatives both within EU as well as beyond EU that are looking for a flexible, user-friendly platform to monitor and disseminate their developments is already under discussion.

The European e-Infrastructures Observatory needs to provide a unified, integrated visualization tool that will be easy to generalize and create a standalone application. In this respect, the developed European e-Infrastructures Observatory should gradually migrate beyond just benchmarking services but also include other e-Infrastructure one-stop-shop services. This is already taking place with the enhanced services that are currently being deployed such as the Pan-European Networks and Stakeholders mapping, the Chronology, etc.

Overall, in order to achieve the long-term exploitation/sustainability goal, the European e-Infrastructures Observatory will make sure it evolves to an indispensable online dissemination and collaboration platform for e-Infrastructures, meeting the needs of the scientific communities, the e-Infrastructures stakeholders and the public at large.

Eventually the European e-Infrastructures Observatory, besides its core function of providing a database of ‘raw data’ and indicators and graphical representations, could become with the addition of the above tools –and after appropriate time for maturity- a kind of an “e-Infrastructure library”, a registry of e-Infrastructure-related components and an online collaboration platform where the e-Infrastructure stakeholders are mapped and profiled, thus allowing “speed-dating” between those who provide the e-Infrastructure resources and those who are using them: the targeted scientific communities and the public-at-large.

5. Conclusions

e-Infrastructures is an important growth area not only for Europe but for the entire world, enabling research and scientific communities to work together by supporting individual collaborative initiatives in ways that were unthinkable a few years ago.

The availability of state-of-the-art European e-Infrastructures to the research community is part of the Europe 2020 vision and the European Commission’s growth strategy for a smart, sustainable and inclusive economy. The Grid infrastructure is a core building block of the European e-Infrastructures and thus is anticipated to majorly contribute to Europe’s Digital Agenda strategic priorities.

Following the continuous investments in e-Infrastructures, there is a clear need to assess European Commission investments in the area of Grid and other electronic Infrastructures and to use this as feedback when developing new European Union policies in this area. Although

economic evaluation and modelling in other areas of ICT have been on-going for some time now, quantitative evaluation of eInfrastructures' development and impact is still at early stages.

Defining, measuring, quantifying and, in general, tackling policy and programme evaluation is a multi-faceted challenge. Evaluation should not be a one-time event, but rather a continuous, on-going process that evolves -becoming better, more thorough and more accurate in its reflections- alongside the very subject that it evaluates. Furthermore, there is no "one-size-fits-all" methodology for performing evaluation and many variations can be applied.

Within this framework, the European e-Infrastructures Observatory responds to the e-Infrastructures community's need for a tool to support progress monitoring and evaluation of related e-Infrastructures initiatives and programmes. Through the collection and utilisation of transparent and reliable data, it monitors e-Infrastructures development and communicates all findings to related stakeholders and to the public-at-large, in a seamless and impartial way.

Such a facility can be of great interest and value to EGI.eu and its stakeholders, as a proper dissemination framework of the respective grid developments at national and pan-European level, as well as a comprehensive and structured statistical framework for measuring the status of development of Grids and other e-Infrastructures across the European Union and beyond, helping stakeholders and policymakers to monitor, analyse and measure economic, social, research and education benefits.

References

- [1] Ramboll Management-Matrix-Eureval Consortium, *Community Support for Research Infrastructures in FP6: evaluation of pertinence and impact*, Study report commissioned by DG Research and DG Information Society and Media, September 2009, retrieved at: http://ec.europa.eu/research/infrastructures/publications_en.html.
- [2] *Role of Research Infrastructures for a Competitive Knowledge Economy*, in proceedings of the *European Strategy Forum on Research Infrastructures Workshop*, DG Research, Brussels, June 2009, retrieved at: http://ec.europa.eu/research/infrastructures/impact_studies_en.html.
- [3] *Work Programme 2010, Capacities, Part 1-Research Infrastructures*, European Commission C(2009)5905 of 29 July 2009, retrieved at: ftp://ftp.cordis.europa.eu/pub/fp7/docs/wp/capacities/infrastructures/n_wp_201001_en.pdf.
- [4] *Responding to Strategic Needs: Reinforcing the use of evaluation: Communication to the Commission from Ms. Grybauskaitė in agreement with the President*, SEC(2007)213, Brussels, February 2007, last retrieved at: http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/sec_2007_0213_en.pdf.
- [5] *8th e-Infrastructure Concertation Meeting*, last viewed at: <http://www.e-sciencetalk.org/e-concertation/>
- [6] *European eInfrastructure Success Stories: The eInfrastructure developments and future perspectives told by the success stories of the European funded projects*, European Commission Information Society and Media Directorate-General, 2008.
- [7] *Description of Work* (2010), e.nventory project, in *Annex I grant agreement 261554*.
- [8] *European e-Infrastructures Observatory website*, www.eInfrastructuresobservatory.eu.
- [9] *World Telecommunication/ICT indicators*, *International Telecommunication Union*, last viewed at: <http://www.itu.int/ITU-D/ict/publications/world/world.html>
- [10] *TERENA Compendium of National Research and Education Networks in Europe*, retrieved at: <http://www.terena.org/activities/compendium/>