



e-Infrastructure Reflection Group supporting the development of the pan-European e- Infrastructure

Wednesday 17th February 2010 @ Public launch of the DC-NET
project, Rome

Leif Laaksonen (e-IRG Chair)
CSC - IT Center for Science Ltd.

e-IRG
e-Infrastructure
Reflection Group

www.e-irg.eu

Rita Colwell: A Global Thirst for Safe Water: The Case of Cholera (Southeast Asia)

- “We explore the problem on different scales. We study the relationship among bacteria, its copepod host, and many other ecological and social factors. On a microscopic level, we look at molecular factors related to the toxin genes in vibrios.”
- “Our challenge today is to apply all our tools of vision, from genomics to computers to satellites, to embrace and understand complex systems.”
- “Although the explanation of cholera outbreaks is global in scale and includes many factors, the solutions can be surprisingly simple--and these can be implemented on a very local scale.”
- “Our most recent results underscore the absolute importance of including social scientists in a project.”

Trends driving the research

- System level science
 - The integration of diverse sources of knowledge about the constituent parts of a complex system with the goal of obtaining an understanding of the system's properties as a whole [Ian Foster]
- Multidisciplinary research
 - Each discipline can solve only part of a problem
 - Collaboration between different research groups
 - Distributed across states, countries, continents
- Research driven by (distributed) data
 - Data explosion, both in volume and complexity
 - Simulation and experiment combined
 - Exploring data-sets with no up-front hypothesis
- Research carried out using simulation and modelling
 - HPC and Grid computing together with high speed networks enable totally new visions in simulations of complex phenomenon

Services for scientists, engineers, scholars, and citizens

- The ICT infrastructure for science is called e-Infrastructure and is the elementary building block of the e-Science.
- The term Cyberinfrastructure was used by a US National Science Foundation (NSF) blue-ribbon committee in 2003 in response to the question: how can NSF, as the nation's premier agency funding basic research, remove existing barriers to the rapid evolution of high performance computing, making it truly usable by all the nation's **scientists, engineers, scholars, and citizens**? Cyberinfrastructure is also called e-Science (<http://en.wikipedia.org/wiki/Cyberinfrastructure>)
- e-Infrastructure is an environment where research resources (hardware, software and content) can be readily shared and accessed wherever this is necessary to promote better and more effective research' (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:EN:PDF>)
- The term e-Infrastructure refers to the new research environment in which all researchers - whether working in the context of their home institutions or in national or multinational scientific initiatives - have shared access to unique or distributed scientific facilities (including data, instruments, computing and communications), regardless of their type and location in the world.
- Intensified measures have to be taken to be able to explain to the society why the e-Infrastructure matters and how the e-Infrastructure and research relate to each other.

The importance and value of data in research

(P. Arzberger et al. Data Science Journal 2004)

- Publicly funded research data should be openly available to the maximum extent possible.
 - Data from publicly funded research are a public good produced in the public interest.
 - Both the data from publicly funded research and research itself have strong public good characteristics that support their open availability to the public, and especially to other researchers.
 - Factual data are central to the scientific research process.
 - The production and open dissemination use of factual data are essential attributes of, and inputs to, modern systems of scientific research and technological innovation. Recognizing the role of digital data as fundamental to the value chain of science, technology, and innovation will enable an optimum return on public investments.
 - Data access and sharing issues are international in scope.
 - To more fully exploit the possibilities of global digital networks, and to capture their benefits for the global community, policy issues concerning access to and sharing of publicly funded scientific research data must be addressed, not only at the institutional and national levels, but also at the international level.

e-Infrastructure domains providing functions and services

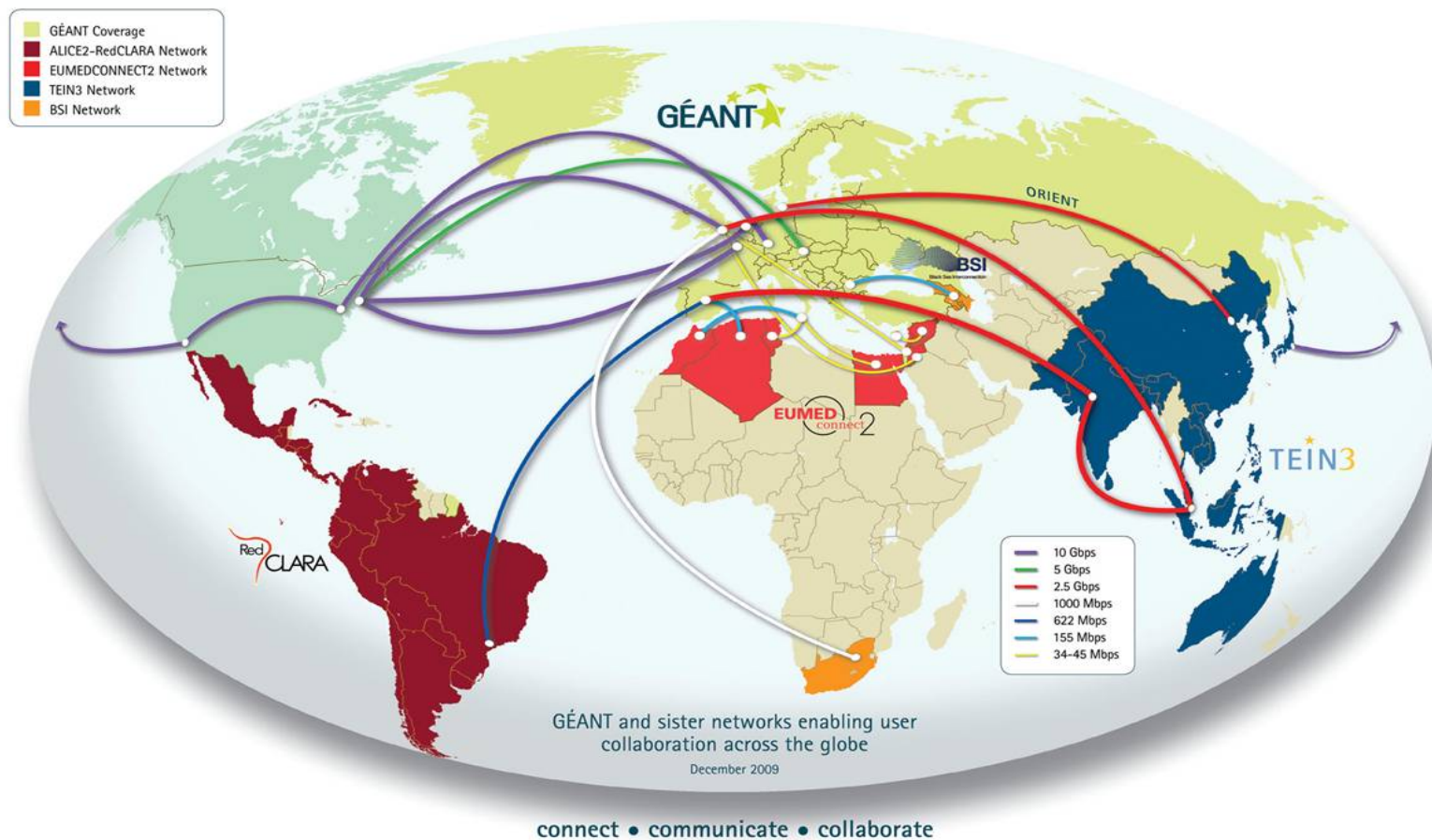


- GÉANT is the world's largest multigigabit communication network dedicated to research and education.
- e-Science Grids respond to the requirements of the most demanding scientific disciplines (e.g. high-energy physics, bioinformatics) to share and combine the power of computers and sophisticated, often unique scientific instruments.
- The scientific data domain tackles the accelerated and uncontrolled proliferation of data elementary of the scientific discovery process.
- Supercomputing e-Infrastructures address the data-intensive and complex challenges of providing modern science with the new computing and simulation capabilities.
- Global Virtual Research Communities, anticipating the advent of research 2.0 paradigms, opens new perspectives for cross-border multi-disciplinary collaboration among research communities.

ICT INFRASTRUCTURES FOR e-SCIENCE
Communication from the Commission 5.3.2009

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0108:FIN:EN:PDF>

GÉANT global connectivity



GÉANT global connectivity – November 2009

Source: Dante

Relevant policy bodies in Europe

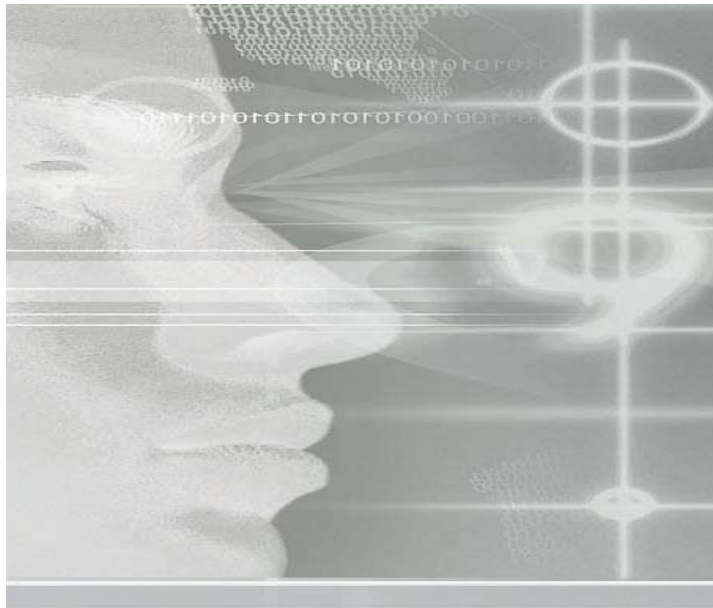


- The European Strategy Forum on Research Infrastructures - ESFRI brings together representatives of EU Member States and Associated States, appointed by Ministers in charge of Research, and one representative of the European Commission.
- The role of ESFRI is to support a coherent approach to policy-making on research infrastructures in Europe, and to act as an incubator for international negotiations about concrete initiatives.
- <http://cordis.europa.eu/esfri/>
- The e-Infrastructure Reflection Group (e-IRG) was founded to define and recommend best practices for the (pan-) European distributed e-Infrastructure. Representatives appointed by Ministers in EU Member States, Associated States to the EU Research Framework Programme and the European Commission.
- The main objective of the e-Infrastructure initiative is to support the creation of a framework (political, technological and administrative) for the easy and cost-effective shared use of distributed electronic resources across Europe.
- <http://www.e-irg.eu/>

Mission and Vision of e-IRG

e-IRG Meeting in Lugano (Switzerland) in 2008:

The e-IRG mission is to pave the way towards a general-purpose European e-Infrastructure.



The vision for the future is an open e-Infrastructure enabling flexible cooperation and optimal use of all electronically available resources.

Roadmap to an ESFRI Research Infrastructure eco-system (update 2008)

*ESFRI addressing
fields of Research and
major research
challenges*

**SOCIAL SCIENCES &
HUMANITIES**

ENVIRONMENTAL SCIENCES

ENERGY

**BIOMEDICAL AND LIFE
SCIENCES**

**MATERIALS AND ANALYTICAL
FACILITIES**

**PHYSICAL SCIENCES AND
ENGINEERING**

**e-IRG building the
e-Science**

E-INFRASTRUCTURE

Crossing the boundaries of science with collaborative computing

Resources/services (capability & capacity computing, sensors, data)

Middleware and organization

Networking Infrastructure

e-IRG aims at global challenges with a high societal impact

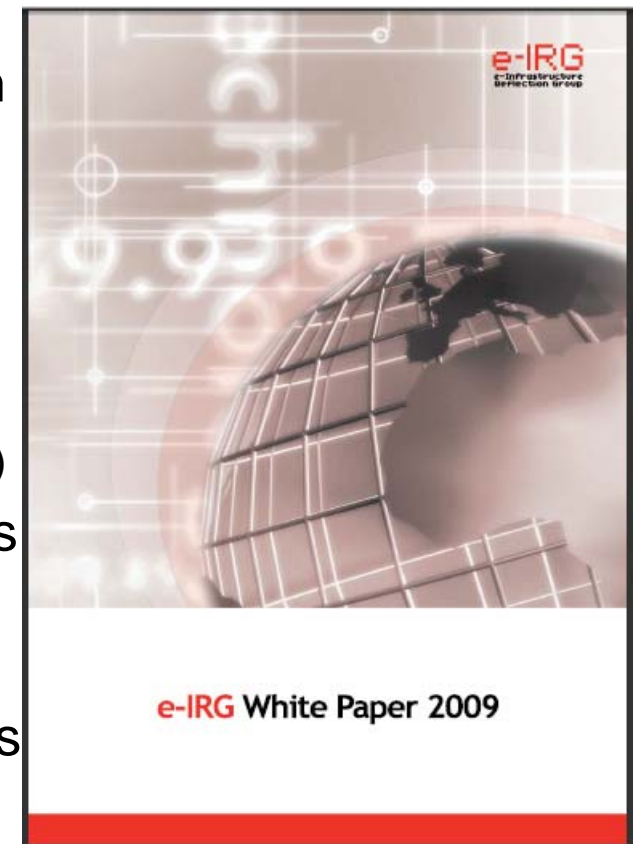


- Targeting audiences:
 - Policy makers on the governmental and inter-governmental levels, dealing with funding, privacy and other issues that are becoming more and more crucial due to the broader uptake of the e-Infrastructure-enabled technologies and processes.
 - Service-providers on the e-Infrastructure domain, such as organizations operating research networks or computing centres – or projects that build on this base in order to provide higher-level multidisciplinary services.
 - Existing and new user communities, looking for a broad overview of the capacities and capabilities that the current and near-future e-Infrastructure can provide.

e-IRG White paper

- One of e-IRG's main publications dealing with e-Infrastructure technologies and projects **requiring policy-level actions**
- A collection of independent and self-sustained topics from different e-Infrastructure areas; Examples:
 - **Grid and Cloud computing**
 - A “holistic” approach to **Security** (seen “as a whole”)
 - **Sustainability** of the computing related e-Infrastructure
 - Service-centric e-Infrastructures through **Virtualization**
 - **Remote instrumentation** (and its integration with e-Infrastructures)
- Topics selected and content produced by e-IRG delegates and experts in a multi-stage consultation process
- Studying technological, economical, societal and policy developments and limitations requiring appropriate actions
- Targeting different audiences:
 - Policy makers, e-Infrastructure providers and users

<http://www.e-irg.eu/> -->
Publications



Data Management Task Force report 2009

- A collaborative effort between e-IRG and ESFRI carried out during 2009.
- The report is divided into three parts:
 - a survey of existing data management initiatives;
 - metadata and quality;
 - interoperability issues in data management.
- The report ends with conclusions of the study and a set of proposed recommendations for further analysis and discussion by the e-IRG.
- The findings will also be presented to the e-IRG and to ESFRI to create a final set of recommendations endorsed by the two bodies.
- Finalized and endorsed by the e-IRG and ESFRI during 2009!

http://www.e-irg.eu/images/stories/publ/task_force_reports/dmtfjointreport.pdf

Roadmap goal, method & timing

- Coherent e-Infrastructure scenarios to support policy formation (2020 and beyond)
 - Opportunities - technology, user communities, use cases,...
 - Service provision: organisational and policy development
 - Requirements: science, society as a whole
- Iterative consultation
 - e-IRG delegates + appointed ESFRI experts, external experts, e-IRG community, general public
- “Discussion opener” for public consultation
 - Final draft in consultation was open until end of January 2010
 - e-IRG recommendations 2010
- Final version is now being processed
 - Public release (hopefully) by the end of February or early March!

Challenges in the e-Infrastructure deployment

- Petascale computing is driving a new trend - how will the scientific software evolution keep up to program hundreds of thousands of processors?
- The development and maintenance of scientific software is an essential, integral component of e-Infrastructure to support research. Software development and life cycle management needs improvement and support.
- Education and training is urgently required to be able to utilize the benefits from the investments in e-infrastructure. A thorough dissemination of already available e-infrastructure knowledge to a much wider workforce and potential user community is required.
- Challenges in having (open) access to data and information resources.
- e-Infrastructure provision must be directed by the needs of the research community and be based on its requirements to carry out major global research efforts.

Sustainability of e-Infrastructures

e-IRG recommendations by the task force 2006:

1. Governments and the Commission should develop policies and mechanisms to encourage increased investment in a more coherent and interoperable way across Europe.
2. The existing e-Infrastructure projects must be superseded by integrated sustainable services at national and European levels.
3. e-Infrastructures must be application-neutral and open to all user communities and resource providers. National funding agencies should be encouraged to fund multi-disciplinary and inclusive infrastructures rather than disciplinary-specific alternatives
4. e-Infrastructures must inter-operate and adopt international standard services and protocols in order to qualify for funding
5. The Commission should, within the seventh Framework Programme, develop a pan-European e-Infrastructure which explicitly encourages the further integration of national e-Infrastructure initiatives.

e-IRG Knowledge Base

- Contains information on European e-Infrastructures
 - HPC, grids and storage resources for science and research
 - National and European policies and policy organizations
 - National and European research networks
 - Projects and initiatives
 - Funding programmes
 - Contact points
- Available both to the e-IRG members and the general public via website (<http://knowledgebase.e-irg.eu>)
- Implemented using standards: Topic Maps and XSLT tools

Summary/Conclusions

- No geographic region has a monopoly on intelligence or creativity.
- Computing and modeling are driving a new trend - how will the scientific software evolution keep up to the programming paradigm? Software lives longer than hardware.
- The development and maintenance of software is an essential, integral component of e-Infrastructure to support research. Software development and life cycle management needs improvement and support.
- Education and training is urgently required to be able to utilize the benefits from the investments in e-infrastructure. A thorough dissemination of already available e-infrastructure knowledge to a much wider workforce and potential user community is required.
- Challenges in having (open) access to data and information resources.
- e-Infrastructure provision must be directed by the needs of the research community and be based on its requirements to carry out major global research efforts.

Thank you!

e-IRG secretariat

**c/o CSC - IT Center for Science Ltd.
(Keilaranta 14)**

P.O. Box 405, FI-02101 Espoo, Finland

Phone: +358(0)9 457 2281

www.e-irg.eu

www.e-irg.eu/support

secretariat@e-irg.eu

