Professor Malcolm Atkinson, NeSC
Elizabeth Vander Meer, NeSC

November, 2007
Updated: April 7, 2008

Policy for Supporting Grid and e-Science Education and Training

Status of This Document

This document provides information to the Grid community on policies for supporting Grid Education and Training. It does not define any standards or technical recommendations. Distribution is unlimited. This is a draft version of the document, it has not yet been submitted for public comment, but is still being developed within the Education and Training Community Group (ET-CG).

Copyright Notice

Report Purpose

This document is a call to action, identifying issues and proposing a strategy in order to support and make progress in grid and e-Science education and training. Inevitably, it is neither complete nor definitive. The intention is that it will seed much greater efforts to further develop the understanding of requirements, to better characterise challenges and to propose specific strategies, curricula and collaborative efforts for international adoption. The ET-CG is already fostering other more specific work and documents[1] that form elements of that development.

Abstract

The development of e-Infrastructure, of which grid computing is a fundamental element, will have major economic and social benefits. Online and financial businesses already successfully use grid computing technologies, for instance. New research methods and technologies generate large data sets that need to be shared in order to ensure continued social and scientific research and innovation. Distributed computing can provide an environment for coping with these large data sets and for sharing data across regions. An investment in educating people in grid computing and e-Science, then, is an investment that will strengthen our economies and societies. In order to deliver e-Infrastructure education internationally, we must develop a policy framework that will ensure shared responsibility and equivalent training in the field. This document introduces the current challenges for grid and e-Science education and training and presents opportunities and existing structures for education and training, as a starting point for further work. It then proposes strategies and policies to provide a supportive framework for e-Infrastructure education and training.

Acknowledgements

The development of the OGF Education and Training Policy Document was supported by the ICEAGE Project and EGEE.

Contents

1. Introduction – why invest in e-Infrastructure education and training?

The United States National Science Foundation (NSF) presents a definition of ‘cyberinfrastructure’ which can be used to describe the various components of e-Infrastructure (the words are interchangeable)[2]:

> Computing systems, data, information resources, networking, digitally enabled sensors, instruments, virtual organizations, and observatories, along with an interoperable suite of software services and tools. This technology is complemented by the interdisciplinary teams of professionals that are responsible for its development, deployment and its use in transformative approaches to scientific and engineering discovery and learning.

It is precisely these teams of professionals that are the focus of the OGF Education and Training Community Group (ET-CG). Investments in e-Infrastructure require adequate investments in e-Infrastructure education in order to allow for the full development and utilisation of these technologies.

We can, therefore, identify four primary reasons why it is vital to develop policies to support e-Infrastructure education and training:

1) a skills and knowledge shortage in business, government, academia and society
2) optimisation of the use of e-Infrastructure
3) benefits for industry and academia
4) the transition to knowledge-based economies

Skills and Knowledge Shortage

First, we are facing a wide-ranging crisis due to a shortage of e-Infrastructure related skills and knowledge in business, government, academia and society. A 2007 review of IT skills and careers in the UK revealed that skills shortages and skills gaps still plague the field of computer science, and this has knock-on effects in other subject areas and sectors. Skills required to use virtualisation technology, for instance, are sorely lacking to the point that more than half of UK businesses cannot take advantage of this technology.[3] But the crisis is not only felt in the UK. In 2005, IT companies pointed to a European-wide “skills crisis as a shortage of computer graduates and a retiring technical workforce threaten to bite IT departments by 2006.”[4] The European Commission ICT (Information and Communications Technology) Skills Monitoring Group compiled specific data on skills shortages across the EU Member States and in the United States in a synthesis report which provides details of the crisis.[5]

The European e-Skills Forum, established by the European Commission, reported on the impacts and reasons for the current situation[6]:

> E-skills shortages, gaps and mismatches threaten productivity development within both the ICT industry and the user sectors and this combined effect on European competitiveness is likely to be significant…Chronic significant shortages of ICT practitioner skills have been endemic in most advanced economies, due largely to the very fast growth of ICT activity in comparison with the relatively low supply of new entrants with a relevant *tertiary education qualification*.

In the United States, similar realities were identified at the close of 2007: "there is a distinct shortage of certain IT [skills], and that shortage seems to be growing."[7] In the Asian Pacific region, countries such as Australia and Thailand also struggle with a skills shortage in the IT, or ICT, sector. The Australian Computer Society has highlighted the problem within Australia[8]:

> The Australian ICT industry is in danger of stagnating…growth of industry is being hampered by key skills shortages, falling telecommunications employment and a downtrend in investment in research and development (R&D)…lack of investment in R&D and information knowledge creation is holding Australia's ICT industry back.

In Thailand, a similar situation is playing out which has detrimental results for the country's economy:

> Thailand suffers from having an insufficient number of skilled workers in the high-technology sector…Today many foreign companies look to Thailand, discover that there is a shortage of ICT talent, even if they are prepared to pay high salaries, and so choose alternative countries in which to base their operations.

Concerned computer scientists in the UK have particularly pinpointed a general lack of expertise in grid computing, explaining that[9]:

> Grid may be the liberal arts of computing. It requires knowledge about many IT disciplines, a flexible management approach and acceptance of new ideas. *But resumes boasting grid-specific skills and accomplishments remain rare. Grid is not widely taught, and IT workers with hands-on experience in this young field are tough to find.*

Education in e-Infrastructure is broadly deficient in its current state, as evidenced by the distinct deficits in skills and knowledge noted above. This crisis cuts across regions and sectors, as e-Infrastructure computing technology proves to be a ubiquitous enabler. If the crisis is addressed, we will find ourselves in a win, win, win situation, in which students gain employability, employers gain skilled staff and educators gain a market.

Optimising the Use of e-Infrastructure

The second reason to develop education and training policy relates to the first. e-Infrastructure technologies such as grid computing involve the potential risk of poor return on investment if measures to support usage of the infrastructures are not put in place. A compelling example, which applies to any research infrastructure, is that it takes years of training to get the best out of facilities. Gaining the best from e-Infrastructure is not simply running the most jobs or the largest volume of data. Not even is it about engaging most users, though these are all important factors. The crucial measure of success is the extent to which it accelerates and enables innovation, generates wealth and promotes well-being. The complexity, novelty and changing nature of e-Infrastructure means that there is a high risk of under-utilisation, or non-optimal exploitation without adequate investment in education and training. The investment in e-Infrastructure to date has provided a pervasive and dependable platform on which a relatively small proportion of experts can demonstrate the high value of the research and innovation it enables. Today's challenge is to strengthen this platform so that the realisation of these benefits of e-Infrastructure *becomes routine*, that is, *any* researcher in *any* discipline routinely uses the resources e-Infrastructure provides as fluently as an artist uses a brush or an engineer uses differential equations. This requires two concurrent and coordinated advances:

1) The educational progress identified in this document, and
2) The steady improvement in the facilities, tools and ease of use of the pervasive e-Infrastructure.

At present, the second branch of this strategic requirement is limited by the lack of sufficient skills across a sufficiently broad spectrum of society and academic disciplines to deliver the advances.

Benefits to Industry and Academia

Both industry and academia benefit from e-Infrastructure, or grid computing, outputs or applications, another key motivation for developing policies to progress grid education. Use of e-Infrastructure has already become integral to finance and online businesses, a primary reason for their economic success. In finance, grid computing can solve problems associated with large and complex computations. Data centres at online companies such as Google and Amazon use forms of grid computing to manage the vast number of searches requested by users on a daily basis worldwide. Advances in scientific and other knowledge as well as new technologies have also generated vast amounts of data that require proper management. Phil Wadler, Professor of Theoretical Computer Science at University of Edinburgh, observes that[10]:

Computing has become a fundamental tool in all research disciplines, which often proceed by compiling and managing large databases and/or exploiting computer models and simulations (a topic sometimes called e-Science).

Today's research into social and scientific issues and problems is not only more data intensive, but has become increasingly more collaborative, which often involves the international sharing of data. Education and training in the use of e-Infrastructure prepares students to use grid computing and other systems and these systems facilitate better management of data and collaboration.

Enabling Knowledge Economies

Finally, countries around the world are transitioning to knowledge-based economies, which rely on the education of citizens in the latest ICT and research methods (etc.). OECD and World Bank country studies have confirmed an obvious correlation between investment in education and quality of life and GDP.[11] There are economic benefits to educating citizens and particularly in preparing them, through education, for the current social context in which we see evidence of the use of computing technologies across academic disciplines and generally in our daily lives. Basic ICT infrastructures now exist in a majority of universities worldwide. The European Commission (EC), for instance, has recognised ICT as key to a knowledge-based economy and social cohesion, and so it must have a place in education and training.[12] Individuals can make the best judgements and make contributions to the knowledge-based economy if they are equipped with the proper skills to exploit existing and rapidly developing technologies. e-Infrastructure is one such technology which can provide the tools to allow countries to "become better at producing knowledge through research, diffusing it through education and applying it through innovation"[13] (integrating the knowledge triangle), in order to successfully compete in the global knowledge-based economy.

Summary

It is clear that greater investment must be made in e-Infrastructure education so that a skilled workforce exists to use and further develop e-Infrastructure technologies throughout the world. Without education and training that targets both students in computer science, those individuals who need in-depth operational knowledge of e-Infrastructure, and students in other disciplines, who must know how to use e-Infrastructure to enhance their research or work capabilities, countries around the world will flounder in their attempts to become players in the knowledge-based economy.

As a means of confronting and correcting the skills and knowledge shortage apparent in e-Science and grid computing technology, this OGF Report argues for further investment in e-Infrastructure education and training; it envisions the embedding of education and training into normal academic training throughout the world. This Report presents a list of motivations that justify this vision, after introducing relevant vocabulary. It provides a picture of the current state of grid and e-Science education in particular. It highlights challenges, those areas that need improvement and development, and identifies opportunities, any existing methods and tools that can be used and built upon. The Report concludes by proposing strategies and policies that delineate the vision for continued coordinated international growth of e-Infrastructure education, thus allowing for the full exploitation of e-Infrastructure technologies.

We specifically need to:

1) Invest in education in appropriate computational thinking or digital-systems judgement in every scientific, medical, engineering and humanities first degree so that a culture is developed and graduating students are equipped to contribute to the knowledge economy with an appreciation of the potential of e-Infrastructure and rich information sources and well prepared to make competent ethical and socio-economic judgements about their use.
2) Invest in education of specialists via undergraduate courses and Masters courses to develop a critical mass of experts who will innovate both in the provision and exploitation of e-Infrastructures and e-Science methods.
3) Invest in Doctoral and Postdoctoral training programmes that develop intellectual and business leaders and educational leaders who will take forward the development of international research and innovation capacity in this field.

By harmonising and collaborating internationally, each country will benefit, both from economies in the cost of the required innovation in educational provision and in the mobility of the resulting skilled citizens. The harmonisation also leads to a community of experts and leaders who are better equipped for trans-national cooperation in research, innovation and business.

2 Definitions

The full set of definitions used in this document can be found in the OGF ET-CG Glossary of Terms.[14] Those central to the presentation are repeated here.

- **e-Infrastructure –** the term is used to denote the digital equipment, software, services, tools, portals, deployments, operational teams, support services and training that provide data, communication and computational services to researchers, innovators and decision-makers. An e-Infrastructure is usually multi-purpose and has to be a sustained dependable facility which can be used for the duration of the work being done.[11]
- **e-Science –** the invention and application of computer-enabled methods to achieve new, better, faster or more efficient research, innovation, decision support or diagnosis in any discipline. It draws on advances in computing science, computation and digital communications.[15]

- **t-Infrastructure –** e-Infrastructure adapted to the needs of education, trainers and students. Shared t-Infrastructure would be usable by students and teachers internationally, providing easy access to educational exercises running on good emulations of e-Infrastructure.

3. Challenges for grid education

We can find evidence of gaps in current e-Science education, as well as training, which could pose problems for both students and educators teaching the use or provision of this e-Infrastructure. Grid education and training are only one element of the total requirements for education in this field. Further work is planned within the context of OGF to develop a broader agenda. In order to contribute, please see the OGF wiki at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/HomePage.

Certain tools or structures are missing in e-Science education and this could hold back development attempts. Understanding these challenges is important when considering what strategy and policy recommendations to propose. The following challenges have been identified:

3.1 Curricula and textbook development

Key challenges concerning curricula involve the need for concerted coordinated work on its development as well as determining various modes for delivery of curricula. Not enough time has been spent developing and defining curricula for grid and e-Science education. The skills and knowledge developed need to be attractive to industry and academic sectors, since students will be drawn to courses if they are generally assured employment after completion. More time spent on curricula can lead to clarification of the “what” and “how” of teaching grid and e-Science education as well as the drafting of a framework for curricula that can be used internationally.

Content Development – Building on the ICEAGE Curricula Development Workshop

The ACM produces curricula guides for computer science courses and these can be used as a reference.[16] The ICEAGE (OGF-ETTF) Curricula Development Workshop, held in Brussels in February 2008, has resulted in useful collaborations and progress in the area of digital-systems thinking and e-Science education which can also be referenced by the OGF ET-CG (see Appendix A). A framework for curricula has been formulated which calls for uptake across disciplines of an undergraduate course that introduces digital systems thinking, which is akin to Jeannette Wing’s notion of “computational thinking”. Subsequent courses proposed at the Workshop build on this foundation to develop students’ skills in e-Science, to Masters level.

The development of curricula that teaches computational thinking skills has been encouraged and promoted by Wing at Carnegie Mellon University in Pittsburgh, USA. This rallying cry from within the field of Computer Science identifies the broad relevance of computational thinking skills to all disciplines[17]:

> Computational thinking is a fundamental skill for everyone, not just computer scientists. To reading, writing, and arithmetic, we should add computational thinking to every child's analytical ability…Computational thinking involves solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science.

Wing's work has led to the creation of the Center for Computational Thinking at Carnegie Mellon, which calls for interdisciplinary uptake of computational thinking skills. Digital-systems thinking will continue to be developed through OGF associated workshops; at OGF22, the ET-CG proposed organising a second Curricula Development Workshop, to be held in Barcelona just prior to or as a session at OGF23 (June 2008).

Teaching the Use of Models

Grids, web services and other forms of distributed computing allow the pooling of resources, the integration of models and the management and analysis of large data collections. To equip students with the ability to use models appropriately and to have good judgement about the validity and interpretation of results they need experience with models appropriate to their discipline. These may be numerical, stochastic, Bayesian, process, statistical or logical models. Additionally, different disciplines have different tools, such as Matlab, that are used for accessing models, organising parameter sweeps and analysing results. The academic curriculum should give the students relevant experience, preferably using examples related to their discipline and academic maturity, of choosing models, planning their use, conducting *in silico* experiments and interpreting results.

Data Management Skills

Many disciplines depend on increasing volumes of shared data in public or proprietary data repositories. An archetypal example arises in earth systems disciplines which are concerned with predicting climate change and mitigating its impact. Examples of the scale and complexity of this data can be seen in the European INSPIRE Project.[18] The curriculum has to teach students how to find and understand the data relevant to a problem in their field. They need to be able to assess the fidelity and temporal validity of such data, to conduct analyses using that data and interpret the results. They also use different data collections as resources, such as those held at the National Institute of Health, at the European Bioinformatics Institute, NASA and the European Space Agency, which they may need to access, compost data from, analyse those compositions and visualise results. The curriculum has to contain relevant examples and be supported by the resources (student-accessible data, computation and tools) that will enable students to develop the relevant understanding and skills.

Learning About Data Collection Processes

In many subjects, digital data is collected via a variety of instruments: telescopes, satellites, sensor networks, observational buoys, medical images, social surveys, etc. In socio-economic, political, epidemiological and ethnographic research, the data may be produced as a side-effect of people’s daily activities. Still other data is generated by collaborating communities subscribing to compendia of observations and annotations. Students require an appreciation of the data collection processes and the ways in which data may be post-processed to generate derivative information, to normalise and standardise to deal with equipment and observing variation and so on. The topics taught must again be relevant to the given discipline and develop judgement as to the interpretation of such data.

Forecasting and Data Mining Skills

More advanced students, e.g. those engaged in forecasting the course of exceptional environmental events (floods, hurricanes tornados, eruptions, earthquakes, tsunamis, etc.) require understanding of the challenges of coupling observation and modelling, and of meeting time constraints in delivering results. This may lead on to all of the issues that arise in planning and coordinating emergency response. Data mining is widely used in some disciplines. Students in these disciplines require an understanding of the forms of data mining and how these may be used with distributed resources and the interpretation of the results they produce.

Knowledge of Ethical Issues

All students need to develop a professional understanding of the ethics of information systems as policy may one day depend on the quality of their advice. They should have an understanding of privacy issues, encryption techniques and security methods. Here again practical and valid examples relevant to the discipline and academic maturity of the students are necessary. It would be good if in the longer term this could build on a core of general knowledge and developed judgement that could be assumed for the intake of e-Science courses.

Proposed Student Categories – Professional Development of Engineers

The majority of students in tertiary education will be in disciplines in which the primary educational goal is to better enable them to be expert users of e-Infrastructure. There are also students, in a range of computing science, informatics, computational science, engineering, mathematics and statistics, who may have careers that contribute to the relevant technologies and systems. For this cohort, the following breakdown of student category and educational needs for professional development of engineers responsible for delivering e-Infrastructure and new software systems and tools that exploit it was identified at the 2nd ICEAGE Forum[19]:

Computer scientists and software engineers—theoretical foundations of distributed computation and insights into engineering trade-offs and current implementation strategies.

Application developers and users—functional and pragmatic presentation of capabilities, an understanding of performance and cost trade-offs and illustrations tuned to their disciplines.

System engineers and managers—criteria to assess and select technologies, need to understand operational trade-offs and failure modes, and need to be able to undertake resource planning.

Modes of Delivering Curricula

Content of curricula has been considered above, but we also must consider modes of delivery. Multiple modes of delivering distributed computing education would be required not only to address the issue of fluidity of the technological landscape (highlighted in a subsequent section) but in order for that education to have wider appeal and relevance and thus greater uptake. Different target audiences would require the presentation of different principles, concepts, and examples, so that the mode of delivery and curriculum are geared towards that audience. Flexible refresher courses could update students on new technologies and summer schools could appeal to academics who would not have time to commit to a Masters course in grid computing or e-Science.

Professionalising Grid Computing and e-Science

More time spent on curricula can lead to progress in professionalising a new category of engineers specialising in grid computing and e-Science, to establish professional practices after refining curricula to meet the needs of various types of students. At the 2nd ICEAGE Forum, this need for professionalisation was raised after discussions concerning how to improve on current systems unreliability and failings.[20] Contemporaneous work by the ET-CG addresses aspects of this requirement.[21] Accreditation bodies such as the BCS and UK Engineering Council, and ACM, could play an important role by certifying courses so that students completing these courses have degree credentials that are widely recognised and which are therefore more valuable to potential employers. And, as explained above, curricula development should include a broader view that presents ways to teach digital systems thinking so that associated skills are embedded across disciplines in all professions.

e-Science Textbooks

There currently is a lack of adequate textbooks to support curricula. e-Science educators face the challenge of writing good text books, as do educators in other fields, which require clarity and conciseness so that students can grasp complex ideas and concepts. It takes time to know how to teach distributed computing well "as a whole". You need to know what to teach (what to leave out) and how to teach it (considering method, structure/organisation of material). There is a still greater challenge if you set out to equip students in a cohesive group of disciplines how to take best advantage of e-Infrastructure.

How to Generate Textbooks

One way to generate textbooks would be to set up a fund to pay for selected leaders in the field to devote time to writing (one year, for instance). Another option would involve the pooling of information on specific sites, sharing this information and debating about what and how to teach, coming to consensus and developing (the outline of) a textbook from this, which can be used internationally (translated). Cooperation on the creation of this textbook would lead to improved resources for teaching (and more efficient development of these resources). A strategy would need to be developed to determine how to go about this and in turn, policy would need to be developed regarding pooled information and its use in textbooks. The SURA Grid Technology Cookbook has recently been made available online and could provide a guide in terms of content for future grid computing textbooks, but also in terms of the collaborative efforts involved in its creation.[22]

The majority of students in tertiary education will be in disciplines in which the primary educational goal is to better enable them to be expert users of e-Infrastructure. There are also students, in a range of computing science, informatics, computational science, engineering, mathematics and statistics who may have careers that contribute to the relevant technologies and systems. For this cohort, it is recommended that incentives be developed, e.g. a competition, in conjunction with established editors and publishers, to develop textbooks that serve and help to define the agreed educational goals and curricula.

Ultimately, the normal commercial processes leading to established and progressively improved text books will probably take over the field, but this depends on developing a market of sufficient size. The initial steps described above are needed to build such a market.

3.2 General Expertise in e-Science: grid computing and computer science

We can identify a lack of "general" experts in the field of e-Science, and a shortage of experienced teachers. Development of education would involve the sharing of material, as expertise in certain areas of e-Science is scattered among individuals**.** The challenge would be to create a new approach to managing and sharing teaching materials due to this lack of general experts, in order to advance academic and research communities.

It is necessary to prime and stimulate an incremental international growth in e-Science educational capability. This has already started in countries throughout the world, partly due to the effects of the Information Society Digital Infrastructures programmes[23]. It requires a positive feedback loop of the following form:

1) Research on infrastructure R&D generates experts with knowledge of e-Science
2) Some of those experts' time is then invested in developing curricula, courses and material and in educating a cohort of students.
3) Some of those students enter step 1 with greatly increased skills and knowledge compared with their forerunners and in increased numbers.

Initially step 2 is achieved mainly in doctoral and post-doctoral programmes. To increase the step change in skills, knowledge and capabilities this must now move into the undergraduate programmes.

Computer scientists contributing to the development of e-Infrastructure education will most often be specialists in a particular technology within their field, which can be problematic when attempting to expand education beyond that aspect of computer science or when teaching how methods may be used in a particular discipline. But, computer science need not provide e-Science education across all disciplines. It can, however, provide other disciplines with the basic tools necessary to incorporate grid education into their academic departments, to become a force for sharing materials and allowing access to experts.

3.3 Teaching and the fluidity of the technological landscape

It is difficult to keep up with rapid change in the computing world. Grid technology and associated standards are constantly evolving with new recommendations and software from standards bodies and solution providers.[24] This means that educators have a daunting task, as do students attempting to learn ever-changing material. Grid computing can provide the solution by strengthening collaborations and cooperative networks which can result in better understandings of these changes and rapid international response, leading to advancements across disciplines and an overall increase in competitiveness. An opportunity arises to develop policies and institutions to facilitate fast and fair exchange.

3.4 Disparate educational policies: harmonisation and security

Harmonisation

Pertinent educational policies that already exist in universities and within countries (at national level) are disparate. For example, university grid access policies for students differ from country to country and even within countries; currently in the UK, postgraduates can have access to the National Grid Service (NGS) but project, campus and regional grids can often have a variety of student access policies and this is problematic.[25] There is a need for harmonisation of these policies so that grid computing is introduced (with ease) more broadly within most disciplines. Students and teachers need to be able to reuse skills and experience as they move around the world. There is a need for policy harmonisation or mechanisms to support interoperation**,** since grid computing is generally international. Grid education can be promoted and use of grid computing can be increased through harmonisation of these education policies, for the benefit of users and providers.

Some students will require practical and specific skills, such as the description and submission of computational jobs, the management and movement of files and the coding of programs to execute in and exploit a grid context. Here the widespread adoption of relevant standards, *including in the taught material*, is an obvious step towards harmonization. In the examples just given, the OGF standards, JSDL, GSM, GridFTP and SAGA[26], would probably be the basis for consistent treatment, leading to skill (as well as code) mobility.

Security

Following on from the challenge to harmonise education policies is the challenge of security. Security issues arise as a result of the sharing of resources across institutions and state boundaries, leading to access and use problems.[27] For example, universities issue identity and authority for students to work with their facilities. When students and staff use multi-institution or multi-country facilities some risks of misbehaviour and choice of authority occur. But complex authorisation can inhibit engagement.

In order to move towards policy harmonisation, the conditions of use that would need to be placed on students, home institutions and visited organisations (this division may not be applicable, depending on how grid access and use is determined, but it provides an example of possible tiers of responsibility) and the providers/operators of grid computing services, as well as technical requirements, would have to be clearly defined and communicated. The eduroam infrastructure use policies (including the European eduroam confederation policy) and technical specifications can provide starting points for future work on such requirements and development of harmonised international e-Infrastructure and grid education and training policies.[28]

Students need to be allowed secure and clearly-defined access to and use of resources (what they are allowed to do must be clearly understood) through authorisation structures as they learn and develop knowledge and skills.

3.5 Sharing training infrastructure

The term t-Infrastructure is used to denote the infrastructure that is needed to enable the educational goals to be met, particularly to develop understanding and experience through practical experience. In a sense, it is the e-Science analogue of laboratories in biology. In practice, the t-Infrastructure is the computing equipment, digital communications, software, data and support staff needed to teach a course. The OGF ET-CG has begun to clarify issues surrounding t-Infrastructure provision in the Training Infrastructure Document, which details European experiences with training platforms such as Gilda and Genius and provides world-wide examples including the Open Science Grid and summer school infrastructures. The document explains the need for a standardised permanent training infrastructure, certainly across Europe.[29]

As the discussion of curricula indicated above, there are many topics to be taught, and their presentation has to be adapted to the discipline(s) and maturity of the students. To give the students good practical experience requires much investment to develop or acquire the relevant t-Infrastructure. This is illustrated by a number of examples:

1) <u>Experience of a parameter sweep using a computational model</u>. The software incorporating the model needs to be written, licensed or purchased. This can be best accomplished by pooled efforts across institutions. The data used by the model needs to be set up. This may require selection and simplification to make the task tractable for students. The parameter space to be explored needs to be chosen by the educators for similar reasons. The computational facilities to execute the model runs and collect the results for each student must be provided. This is demanding as (a) the entire cohort will submit their jobs at approximately the same time, and (b) the students require a response within a reasonable time and a low rate of failures or learning is impaired. As classes run at different times in different places, there is a good opportunity to take advantage of pooled resources.

2) <u>Experience of data analysis</u>. Let us say the students are given access to a set of predictions of a hurricane's path and the census and property data of a relevant region and asked to identify areas where the risk times cost is high so that they receive priority for evacuation assistance. Collecting example data of predicted hurricane tracks is probably relatively straightforward, though a single request to the hurricane centre may be much preferable to many requests to the centre from many educators. However, setting up the census and property data is a much more complex task. It requires negotiation over how much information may be presented. It requires transformation to hide the actual data while still presenting a sensible geographic and social situation. It requires adaptation to show all the educational examples but tractability in the expected time for the expected

category of students. The advantage of doing this work once, sharing the cost and re-using it in many institutions and countries is self-evident.

3) Experience of interpreting medical images. As digital scanning methods (e.g. MRI and digital x-ray) increase it is important to educate medical students in their use. The current volumes of data involved can be substantial, as can the computation to render images according to requested viewing parameters. A pooled resource can have several advantages: (a) it shares the collection, cataloguing, anonymisation, ethics negotiation and privacy costs, (b) because it can draw on data from thousands of centres it can have a far more complete collection of rare diseases and rare presentations for a particular imaging technology, (c) because it draws on non-local populations, accidental recognition is very unlikely, and (d) the larger collection may support better atlases and epidemiology.

4) Experience of working in a collaborative multinational and multidisciplinary team. Many research programmes, engineering projects and policy support activities depend today on effective work in such distributed teams supported by the best Computer Supported Collaborative Working, shared computing and telepresence methods. In order that students can be prepared to work in such contexts, they need to undertake projects in their curriculum that simulate relevant aspects of such collaborative working. Setting this up and supporting it require multi-state collaborative action.

3.6 IPR and sharing

A further challenge relating to sharing and trust models involves Intellectual Property Rights (IPR). A framework for sharing in terms of IPR needs to be in place, but so far no models have been widely accepted at international level.[30] In the European context, the 2001 EU Copyright Directive (Directive 2001/29/EC) is an attempt at standardising, or harmonising, copyright law among Member States, keeping in mind certain modern requirements of the information society, and as such it relates to educational materials that would be shared in the case of e-Science (etc).[31] Considering a wider (international) view, the Berne Convention is well-established and addresses the issue of copyright, as does TRIPs, within the World Trade Organisation (WTO) agreements.[32] The World Intellectual Property Organisation (WIPO) also provides frameworks for IPR that might be relevant. But the challenges arising for e-Science and the sharing involved in use of e-Infrastructures are relatively new and still in the process of being unravelled and addressed. This issue is being tackled within the OGF ET-CG.[33] At present, ICEAGE and EGEE repositories provide (contained) educational materials that can be safely used due to such rights issues having been addressed. Rather than copyright, deposit agreements and creative commons licences could provide a model to apply in e-Science education.[34]

3.7 Training-specific challenges and requirements

Training can be distinguished from education in that training is a targeted short-term process to develop specific skills in a certain technical area, whereas education can be seen as an institutionalised long-term process using conceptual models and resulting in development of a culture (but these are by no means discrete categorisations). In order to increase training opportunities in e-Infrastructures, and particularly in grid computing, certain challenges must be addressed, some of which mirror challenges introduced in discussion of education:

- For instance, lack of teachers with appropriate expertise and the problems associated with teaching in the midst of technological change arise in both the areas of education and training. Developing an internationally-recognised certification process which provides teachers with quality training (and credibility) that includes periodic updating of knowledge would be a reasonable response to this challenge.
- The content of training courses on international level, as well as methods of delivery, are currently different, as they are in educational courses, but in the case of training this is often the result of vendor variety (so that each vendor provides training on their product and each product requires unique vendor-specific methods of operation). “Vendors” should be interpreted liberally here, to include projects such as Condor, DEISA, EGEE, Globus and SRB that deliver technology. Definitions of key terms, for instance “security” and “job”, may differ depending on the vendor, based on differences in product.
- Cooperation on development of shared t-Infrastructure would be beneficial in the training arena.

Despite these similarities and overlapping challenges, certain training-specific challenges and requirements can be identified:

- To define the structure of training certifications, considering skills required at each level. Work has already been done within the OGF ET-CG to suggest types of certificates, based on skill sets.[35] Three certificates have been proposed: certified grid technician (CGT), certified grid professional (CGP) and certified grid architect (CGA). To obtain the CGT certification, the trainee must complete a base technician module and one specialisation module; the focus is on practical rather than conceptual skills. The CGP would obtain a certificate after completing a base engineer module (more in-depth than the CGT base module), more than one specialisation module and after developing both practical and conceptual skills. And finally, the proposed CGA is trained to have a high-level view of grid technologies and their deployment, operation and use.
- To convince vendors (industry players) to participate in developing a general training process.

3.8 Impact of standards on education and training

As remarked above, much of the e-Infrastructure and specific tools in use vary from site to site and in many cases are also evolving rapidly. This variation and the rate of change increases the cost of preparing and presenting courses, reduces skill mobility and detracts from the amortisation of costs through shared t-Infrastructure.

Ineluctably as some of the education goes hand in hand with research, it is at the frontier and must endure rapid change as understanding, methods and technology develops. However, for the majority of the education neither the variety nor the rate of change is necessary.

It is important that the education and training community work closely with the standards development organisations to encourage the development and uptake of relevant standards. For example, the EU education and training community should then work in concert with technology providers, e-Infrastructure providers and educational institutions to encourage and accelerate the adoption of relevant standards. Just as the units used in a Physics course work anywhere in Europe so should the terms and methods taught in an e-Science course. And, in a wider context, the context that concerns the OGF ET-CG, these terms and methods should be consistent worldwide.

4. Opportunities and existing structures for education and training

It is important to understand the existing state of grid and e-Science education in order to know what options are out there for educational planners and how to proceed. Research up to now has focused on identifying existing tools and infrastructures within the European context in particular (resulting in the e-IRG ETTF Report), while we are not as familiar with other world regions. In order to provide a more comprehensive view of the *international* state of grid and e-Science education, in order to explore opportunities, we invite you to provide us with information on your region by visiting the OGF wiki at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/HomePage.

The following identified tools and infrastructure, as mentioned, derive primarily from the EU context, but certainly they reflect what is occurring in many world regions in relation to e-Infrastructure education:

4.1 Existing educational machinery – curricula, t-Infrastructure and security

A number of higher education institutions within EU Member States provide Masters courses and summer schools on grid education. Currently, there are Masters courses available in grid computing and related areas throughout the EU. Certain regions within the EU are beginning to coordinate efforts, such as the Nordic countries using NorduGrid. NorduGrid is developed NGIn, an educational project offering students opportunities to study grid computing at postgraduate level. The Nordic Council of Ministers has also formed a Nordic e-Science working group which proposes establishing postgraduate courses in e-Science. This coordinated effort should be watched and reviewed as a possible example for other world regions to follow.[36] But aside from a handful of exceptions, there appear to be few coordinated efforts across universities to work together on provisions for the Masters courses.

Undergraduate courses and summer schools are run by countries including Greece, Portugal, Germany, Italy, Estonia, Finland, and Hungary. ICEAGE has supported three of the summer schools in the ISSGC series and pioneered an online Winter School.[37]

A list of university postgraduate courses, summer schools and online courses has been compiled on the OGF wiki and contains examples from around the world, but the list is far from comprehensive and requires continual updating.[38]

EU Member States have not yet worked together to create a coherent infrastructure, so that there are no shared security networks and IPR (beyond OGF) and curricula are created on an *ad hoc* basis, without backing from accreditation bodies. Masters courses are aimed at research output (producing researchers) when they could also be aimed at industry through accreditation. Member States could develop a shared t-Infrastructure and shared security and IPR frameworks, as could other regions throughout the world. They could ensure that courses are certified by accreditation (industry and professional) bodies.

4.2 NGIs and the EGI – providing infrastructure for education and training

Building infrastructures is expensive, so coordinating by engaging with regional or national grid system providers already operating in different member states throughout the EU would minimise costs. Coordination that allows sharing of knowledge is also beneficial. Most universities do not have access to all experts in the field, so expert knowledge sharing among institutions would increase the EU's overall competitiveness in research and innovation. Coordination can lead to standardisation of core material and attainment criteria for education internationally, so that mobility is facilitated. Development of an international infrastructure would advance the sharing of curricula, qualifications and teaching methods.

Existing National Grid Initiatives (NGIs) and the European Grid Initiative (EGI) could provide foundational infrastructure for grid education in the EU. There are developing NGIs in 37 European countries which could in principle provide infrastructure for grid education.[39] As a single national point of contact for local institutions in each Member State, the NGI could connect all fields involved in grid computing and e-Science, providing the following services: easily available and accessible t-Infrastructure for classroom exercises and teaching, identity management and security and tools/techniques for setting up "grid in a box" systems on demand. Such an infrastructure can provide a model for other regions throughout the world and also lead to development of linked regional infrastructures that together create an *international* infrastructure.

The EGI, currently in its design phase, will help to integrate the NGIs and provide coverage where no NGIs exist (also stimulating development of NGIs in these Member States). The EGI should help in the harmonisation of e-Infrastructure education across Member States through coordination of NGI services such as authentication and security. The EGI Knowledge Base webpage gathers details regarding the importance and current relationship of NGIs to education and training efforts within each European country; this is a first step towards such coordination and should be referenced to advance work in this area.[40] European e-Infrastructure integration, as well as integrations in other world regions, also has to consider HPC and objectives and activities promoted by the Partnership for Advanced Computing in Europe (PRACE).

Virtually every facility planned in the European Strategy Forum for Research Infrastructures Roadmap[41] has a significant requirement for data management, computation and remote control of experiments. In consequence, the e-Infrastructure and associated education and training are of great relevance to the broad plans for research infrastructures- similar requirements will pertain in every country with experimental or observational facilities.

4.3 Embedded e-Infrastructure in national educational operations, plans and policies

There are already examples of the embedding of e-Infrastructure into national education policies in EU Member States, particularly involving security. In Greece, for instance, students receive their student card, email and grid access upon registration, as part of the existing educational security model.[42] Similar networks that can allow students such access are found in China (ChinaGrid CERNET), Japan (Naregi Japanese Research Grid Project), Spain (RedIris, PAPI) and New Zealand (KAREN).[43] We need to identify whether there are models that would allow this to happen elsewhere.

5. Suggested strategies and policies

This OGF information document sets out options to increase engagement with e-Infrastructure technology, and distributed computing in particular, on an international scale. The document has reviewed the current state of grid and e-Science education, presenting related challenges and opportunities. The suggested strategies and policies listed below can support international development of e-Infrastructure education.

Harmonisation Versus an Organic Process

When formulating strategy and policy recommendations, we have to keep two models in mind: one based on harmonisation and one that allows grid education to develop in an organic fashion, which it is currently doing. There are trade-offs. Advantages to harmonisation are skills transfer, mobility, credit transfer, integration, cost savings and shared curriculum development. Advantages to an organic process are diversity, cross fertilisation which can lead to innovation, meeting national and discipline requirements faster and flexibility to better respond to a rapidly changing domain. Both models will find expression in policies and strategies proposed here.

5.1 Strategies

Curricula development –

Encourage and invest in the interdisciplinary and collaborative development of new grid computing and e-Science modules at departmental, institutional and national levels, and provide means for coordination in terms of curricula:

1) The e-IRG ETTF Report proposes establishing a committee/body of leading educators across disciplines to expedite the creation of the curricula goals and principal topics, launched and supported by major conferences highlighting educational priorities and opportunities in the field.
2) Continue meetings in international contexts, such as that in Brussels and at OGF 22 and 23, to further develop understanding of educational goals and curricula.
3) Continue to build a repository of shared experiences and practice in e-Science education (a list of Masters and other courses is being compiled and can be accessed and added to on the OGF wiki at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/HomePage).

Develop a means to pool information, cooperate and provide standards of use for information to produce textbooks and other teaching material for grid education. Options for production of adequate textbooks include:

1) Establishing specific websites and other relevant fora where information for textbook content can be pooled, shared and debated about.
2) Setting up a fund to pay for a selected leader in the field to devote a block of time to writing a textbook.
3) Developing incentives such as competitions, in conjunction with editors and publishers, to produce textbooks which follow agreed educational goals and curricula.

Refer to the SURA Grid Technology Cookbook and network with contributors regarding content and collaborations. See Appendix B for an expanded list of curricula and textbook development resources.

Investigate changes to education already occurring as a result of emerging ICT and changes that could be made.

Certification –

Encourage certification of courses by professional accreditation bodies, build on current harmonisation and cooperation. At the 21st OGF, E&T sessions included discussion of establishing the Grid Professional Institute (GPI). But at OGF22, ET-CG members concluded that existing bodies such as the BCS (for Europe) and the ACM (for the Americas) should manage certification and this could be done internationally through collaboration and cooperation among these existing organisations.

Promote the sharing of resources –

Investigate shared security models, for t-Infrastructure, relating to existing procedures to move towards standardisation by embedding e-Infrastructure in a similar manner in the national education policies of all Member States. It was suggested at the 2nd ICEAGE Forum that a task force should be set up to assess existing tools, their ease of use and suitability, including security issues. Best practice could be determined after exploring current models.[44]

Address challenges concerning the sharing of materials, considering IPR and repository provisions.

Develop relationships –

Look at national and international e-Infrastructure to support education to determine what relationships to develop. Providing stronger links between the ACM, BCS and similar bodies, in relation to certification would be beneficial. Consider relationships between the DEISA, EGI, EGEE and Open Science Grid (OSG).

5.2 Policies

We can identify the need for two kinds of policy in order to establish a framework for shared responsibility and equivalent educational training:

- **Policy for providers of education**. These would be common rules to address issues arising from the sharing of ideas, software and computing resources.
- **Policy for teachers and students**. These would be common rules to address issues arising from equipment use (so students do not crash systems) including conditions of use and mobility and the need for access (to allow continuity of work, for instance, with PhD students).

The OGF ET-CG therefore recommends the development of policies on the following issues:

1) Recommendations as to the level of investment necessary (nationally) in order to provide education in the use of e-Infrastructure.
(Suggestion: at least 50% of the investment that is going into e-Infrastructure provision. While this figure is significant, it is justified due to the crisis we currently face.
Unless there are adequate numbers of people schooled in the creation, use and further development of e-Infrastructure technologies, countries worldwide will fail to fully exploit these vital tools for research and innovation. The consequences of this failure will be felt both economically and socially and result in losses in the knowledge economy. Ensuring an increase in skilled individuals inevitably involves commitment in the form of funding.)

2) Recommendations as to the harmonisation of education in the use of e-Infrastructure.
(Suggestion: persuade professional bodies, e.g. the Royal Society of Chemists and the Institute for Engineering and Technology in the UK, to identify target attainments for their profession and to harmonise in their region)

3) Propose standards for student and teacher identification that would enable access to educational grid facilities and authorization/management of the resources used.
(Suggestion: build on the eduroam protocols to extend them to cover student use of collaboration facilities and multi-site t-Infrastructure)

4) Propose standards for sharing training material and t-Infrastructure between institutions.
(Suggestion: build on creative commons for all educational material and on NGIs and EGI proposals for t-Infrastructure)

5) Establish a system for agreeing standards that accredit workers who design, build, operate and support e-Infrastructure so that qualifications are recognised internationally.
(Suggestion: adapt the proposals developed by the OGF ET-CG working group[45])

6. Future work

To strengthen the ET-CG Education and Training Policy document, future work will include adding details requested from OGF members on the state of education in continents other than Europe, as the document currently has a European bias. Curricula development will be furthered through a proposed Curricula Development Workshop, to run just prior to or at OGF23 in Barcelona (June 2008). The e-IRG ETTF will continue to be a resource to assist in future development of education and training policy.

7. Contributors

Original Members of the OGF ET-CG

Torsten Antoni
Malcolm Atkinson
Purushotham Bangalore
Rudiger Berlich
Kathryn Cassidy
Oscar Corcho
David Fergusson
Leander Frunklin
Scott Lathrop
Marc Ozonne
Beth Plale
Cong Thanh
Kilian Schwarz
Borja Sotamayor
Eero Vainikko

GGF17, OGF18 and 19 Attendees

NAME	AFFILIATION	Meeting
Maclolm Atkinson	NeSC, UK	Tokyo, Washington, Chapel Hill
David Fergusson	NeSC, UK	Washington, Chapel Hill
Stephanie McLean	Renaissance Computing Institute (RENCI)	Washington
Erwin Laure	CERN	Washington
Wolfgang Gentzch		Washington
Alain Roy	Madison University	Washington
Ben Clifford	Globus & University of Chicago	Chapel Hill
Kathryn Cassidy	Trinity College Dublin	Tokyo, Washington, Chapel Hill
Morgane Artacho	NeSC, UK	Chapel Hill
Rüdiger Berlich	Forschungszentrum Karlsruhe	Chapel Hill
Killian Schwarz	GSI, Karlsruhe	Chapel Hill
Steve Brewer	OMII-Europe	Chapel Hill
Hsin-Yen Chen	ASGC Taiwan	Chapel Hill
Jerry Perez	Texas Tech University	Chapel Hill
Lennart Johnsson	University of Houston & KTH	Chapel Hill

Marc-Elian Begin	CERN	Chapel Hill
Charles Bacon	University of Chicago	Chapel Hill
Christos Kanellopoulos	GRNET	Chapel Hill
Dhivakaran Muruganantham	LBNL/ESnet	Chapel Hill
John Mincarelli	Synopsis Ltd	Chapel Hill
Donal Fellows	University of Manchester	Chapel Hill
Andreas Savva	Fujitsu Laboratories Ltd	Chapel Hill
Stephen Newhouse	University of Southampton, UK	Chapel Hill
Alison Clark		Chapel Hill

OGF20 Attendees

NAME	AFFILIATION
Torsten Antoni	Forschungszentrum Karlsruhe
Morgane Artacho	NeSC, Edinburgh
Malcolm Atkinson	e-Science Institute, Edinburgh
Kenny Baird	NCeSS, Manchester
Azadeh Bararsani	PDC
Ruediger Berlich	Forshungszentrum Karlsruhe
Richard Bruin	University of Cambridge
Antal Bulanza	BELNET
Kathryn Cassidy	Trinity College, Dublin
Peter Clarke	NeSC, Edinburgh
Ben Clifford	OSG, University of Chicago
Jim De Roest	University of Washington
Matthew Dovey	JISC
Yehia El Khatib	Lancaster University
Donal Fellows	University of Manchester
David Fergusson	NeSC, Edinburgh
Ian Frame	NIEeS
Peter Halfpenny	NCeSS, University of Manchester
Suhel Hammoud	University of Brunel, West London
Steinar Henden	University of Tromso
Israel Hernandez	University of Edinburgh
Zack Kertcher	University of Chicago
Vladimir Koren'kov	Joint Institute for Nuclear Research
Jill Kowalchuk	Netera Alliance

Simone Lanzarini	CINECA
Marco La Rosa	University of Melbourne
Yang Liu	University of Brunel
Boon Low	NeSC, Edinburgh
Chris Messom	Massey University
Pierrick Micout	CEA DAPNIA
Enric Mitjana	European Commission
Young Song Mun	Soongsil University
Christoph-Erdmann Pfeiler	Forschungszentrum Karlsruhe (FZK)
Frederic Schaer	CEA
Jennifer Schopf	ANL
Bernhard Schott	Platform Computing GmbH
Gergely Sipos	MTA Sztaki
Federico Stagni	INFN Ferrara
Katerina Stamou	PDC
Anthony Stell	University of Glasgow
Tatiana Strizh	Joint Institute for Nuclear Research
Toshihiro Suzuki	Oracle
Igor Tkachev	Joint Institute for Nuclear Research
Frederique Van Till	JISC
Adam Villa	University of New Hampshire
Gian Luca Volpato	RRZN
Gabriel Zaquine	CS SI

OGF21 Attendees

NAME	AFFILIATION
Malcolm Atkinson	e-Science Institute, Edinburgh
Gerd Behrmann	NDGF
Kathryn Cassidy	Trinity College, Dublin
Mathias Dalheimer	Fraunhofer ITWM
Nishadi De Silva	University of Southampton
Donal Fellows	University of Manchester
Wolfgang Gentzsch	D-GRID
Feikje Hielkema	University of Aberdeen
Chris Higgins	EDINA, University of Edinburgh
Jysoo Lee	KISTI, Korea

Shahbaz Memon	Forschungszentrum Juelich
Anitha Ohri	Platform Computing Inc, Canada
Anand Patil	DANTE
Thomas Prokosch	GUP, Johannes Kepler University, Austria
Adriano Rippa	Engineering R&D
Kilian Schwarz	GSI
Daniel Templeton	Sun Microsystems
Elizabeth Vander Meer	NeSC, Edinburgh
Alex Voss	NCeSS, University of Manchester

OGF22 Attendees

NAME	AFFILIATION
Guiseppe Andronico	University of Catania, INFN, Italy
Malcolm Atkinson	e-Science Institute, Edinburgh
Kathryn Cassidy	Trinity College, Dublin
Oscar Corcho	Universidad Politecnica de Madrid, Spain
David Fergusson	NeSC, Edinburgh
Geoffrey Fox	Indiana University, USA
Hiro Kishimoto	Fujitsu
Kamie Kitmitto	University of Manchester
Thomas Prokosch	GUP, Johannes Kepler University, Austria
Daniel Templeton	Sun Microsystems
Elizabeth Vander Meer	NeSC, Edinburgh

8. IPR statement

The OGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the OGF Secretariat.

The OGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the OGF Executive Director.

9. Disclaimer

This document and the information contained herein is provided on an “As is” basis and the OGF disclaims all warranties, express or implied, including but not limited to any warranty that the use of the information herein will not infringe any rights or any implied warranties of merchantability or fitness for a particular purpose.

10. Full copyright notice

Copyright © Open Grid Forum (2006-2008). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the OGF or other organizations, except as needed for the purpose of developing Grid Recommendations in which case the procedures for copyrights defined in the OGF Document process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the OGF or its successors or assignees.

11. References

References appear after appendices as endnotes.

12. Appendices

Appendix A – Report from the ICEAGE Curricula Development Workshop 2008

CURRICULA FOR UNDERGRADUATE AND MASTERS LEVEL COURSES IN e-SCIENCE:
Report from the ICEAGE Curricula Development Workshop
Brussels, 14-15 February 2008

This document describes an urgent social and economic need to:

1) Equip first degree students in all disciplines with a level of skills in digital-systems judgement or computational thinking sufficient to support and progress the knowledge-based economy.

2) Invest in undergraduate and Masters courses to develop experts capable of innovating in the provision and exploitation of e-Infrastructures and e-Science.

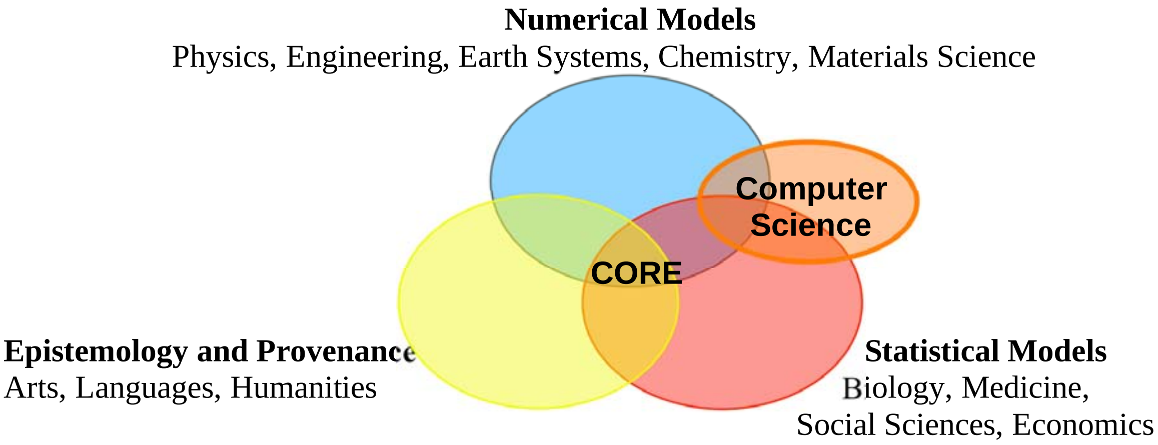
Introduction

The ICEAGE Curricula Development Workshop, held at Scotland Europa in Brussels from 14-15 February 2008, was successful in proposing an initial framework for undergraduate and Masters level courses in digital-systems thinking and e-Science, which can provide the basis for further work. The workshop was co-chaired by Professor Malcolm Atkinson (Director, e-Science Institute, University of Edinburgh) and Dr. David Fergusson (Deputy Director of Training, Outreach and Education, NeSC, University of Edinburgh).

Attendees included Amy Apon (University of Arkansas), Mark Baker (University of Reading), Kenny Baird (NCeSS), Kathryn Cassidy (Trinity College, Dublin), Ben Clifford (OSG, University of Chicago), Joy Davidson (HATII, University of Glasgow), Fotis Georgatos (GRNET), Petar Jandric (NeSC, University of Edinburgh), Fernando Silva (Universidade do Porto) and Elizabeth Vander Meer (NeSC, University of Edinburgh).

The workshop was called by ICEAGE, the OGF ET-CG and e-IRG ETTF and people from all communities were invited to attend. OGF and e-IRG Education and Training policy reports have documented the profound lack of well-developed e-Science curricula at undergraduate and graduate levels, and this deficiency led to initial discussion at OGF21 of running a workshop to provide educators with a forum to begin to address the problem. The Curricula Development Workshop was organized as a result of this call to action.

Developing curricula for e-Science is far from straightforward. Multiple methods and modes of delivery must be considered. Different target audiences would require the presentation of different principles, concepts and examples. It is important to be clear about what students are being targeted, since curricula for computer science students, for instance, would be vastly different from curricula geared towards students in other disciplines, in which numerical models, statistical models or epistemology and provenance may dominate (see Venn Diagram).



Workshop attendees focused on developing a framework for e-Science education across disciplines rather than grid education, which would target students within computer science. It is noted that research and innovation works best if scientific practitioners in the various disciplines operate in close coordination with technical experts in computer science, each engaging with the other's issues. This report does address this matter and stresses the importance of interdisciplinary professional communication, for instance. The need for such engagement becomes clear if a particular project is examined, such as CARMEN (Code Analysis, Repository and Modelling for e-Neuroscience). Both technical experts in computer systems and application scientists are working together to develop a virtual laboratory; in order to refine the way in which the virtual laboratory works once the design has been launched, users must feed back to technicians, there must be constant interchange to ensure successful development. Therefore student experience in more advanced courses should include working in this interface. Practicals can be set up for instance, for geology students to collaborate with computer science students. This kind of collaboration has been given a central role in the following exposition of proposed courses in e-Science.

Core topics and prerequisites for courses were identified during workshop discussions, and it was decided whether courses at undergraduate level should be required or optional. While elements of Stages 1 to 3, which define the undergraduate curriculum, can be found in existing courses, it was decided that these courses would be proposed discretely, rather than worked into the content of courses that already exist. The following report presents the general content suggested for each undergraduate level and the Masters curriculum.

Prerequisites and Educational Goals for Undergraduate Level e-Science Course(s)

The Curricula Development Workshop first focused on developing interdisciplinary content for an undergraduate digital-systems/e-Science curriculum which would be introduced in three stages: e-Working, Basic Methods and Advanced Methods.

STAGE 1: e-Working

The Stage 1 course has been proposed as a requirement across disciplines, available to every student, whereas at present it appears selectively in certain disciplines. There are no hard and fast prerequisites, beyond general university prerequisites, but certain experience would be assumed, including previous use of email, a browser, chat client, word processor, PowerPoint and other software that would be part of a collaborative learning environment.

Stage 1 would be an introductory module imparting students with an understanding of digital-systems thinking, which would provide the basis for further education in e-Science. Educational goals to be achieved by Stage 1 include use of common communication tools in a professional manner, the ability to deal with complex tasks using process thinking (to organize work on tasks as an individual and in groups) and producing results from a team effort that properly reflect contributions and correctly cite material. Tools familiar to the students are used, but Stage 1 stresses the importance of learning how to work together effectively on tasks using these tools, emphasizing collaborative behaviour and provoking students to think about how they are using technologies; use of tools in a shared context involves new skills and presents ethical issues unique to that context. This stage promotes flexible thinking in students and learning through collaborative tasks.

- **Stages & Competencies**
  - Use email, web-search, word-processing, presentation tools with standards in mind, considering the quality of communication (how to use these tools responsibly—ethically, with proper citation and accuracy)
  - Develop team working using the above
  - Use digital-communication tools (work with libraries)

        - To coordinate & develop a deliverable
        - Responsibility, legal, ethical & social issues
        - Security and safety
    - Critical thinking
    - Use of subject-specific digital resources
        - Scientific data and document data
        - Metadata and controlled vocabularies
        - Proper citation and legitimate use
        - IPR
    - Collaborative behaviour
        - Plagiarism and its detection
        - Drawing on strengths & knowledge of team members
        - Strategies for dealing with weaknesses and lack of knowledge

An example of a task assigned to teach these skills could be collaboration on a written report. First students in a group would need to decide who obtains what material for the report (distribution of work, location of resources, how to best access resources). Once they find the resources, they then must consider what should be extracted and how it should be organised, the proper way to cite references, and how to present the report as a team in a coherent manner. To do this, each student has to think about breaking the task down and then reassembling material, thus learning process thinking. The emphasis at Stage 1, as previously stated, would be on working effectively in a collaborative environment using information and communication technologies.

STAGE 2: Basic Methods

Stage 2 will continue to teach competencies introduced in Stage 1, but strands will be tailored to specific disciplines, so that **not all components of the curriculum content listed below would be required core elements**. While aspiring to make the Stage 2 course a requirement, it was decided that it would initially be proposed as optional and with time may naturally become compulsory. Completion of Stage 1 would be the prerequisite for entering Stage 2. As with Stage 1, Stage 2 content remains primarily at a conceptual level and provides students with a mental model of tools for e-Science.

- **Critical thinking (2)***
    - Data curation and management*
        - Data lifecycle, Data Bases, Data models, Metadata, Mark up languages
        - Controlled vocabularies, Ontologies
        - Information entropy
            - *Compression*
    - Complexity
        - Origins of real-world & system complexity
        - Handling complexity
    - Use of models*

        - Validity and domains of applicability
        - Data dependence and interpretation of results
    - Numerical thinking
        - Nature & origins of error
        - Precision, correctness & validation
        - Types of numbers and their behaviour / representation
    - Statistical thinking
        - Sampling and error
        - Uses and abuses of statistics
    - Using multiple data resources
        - Semantic and provenance
- **Responsibility, legal, ethical & social issues (2)***
- **Presentation and interpretation of data**
    - Visualisation
    - Interdisciplinary professional communication*
- **Image analysis**
    - Derivation of information
- **Process thinking**
- **Logical thinking and decidability**
- **Trust: security, privacy and integrity***
    - Risk and impact
    - Implementations, their strengths and weaknesses (not core)
    - Personal behaviour
- **Distributed systems thinking**
    - Digital communication and network services
    - Distributed systems architectures
    - Storage systems and preservation
    - Instrumentation
        - Digital devices, sensors and networks

How much of distributed systems are discussed depends on your audience. For this undergraduate level, it is more important here to recognize that there are models behind commonly used tools such as Facebook and Google. Google, for instance, would provide students with a rough idea of documents on a certain subject, but if they want to find all the references relating to a particular topic or the definitive article on that topic, they would not use this tool (you could contrast Google with a citation index). It would not be necessary to talk about the Bayesian model, but instead focus students attention on the results associated with using such a model (the answers you get). The student would learn that models are tuned to provide certain results, so it is important to recognize what the model is meant to do (what it reveals and what is left out).

Generic Property of Stage 2:

- **Everyone is aware of**
    - the terms and their meaning
    - Where to find experts & more information

- **If we were to get someone from Stage 2**
  - & set up a tool for them
  - Then they can quickly learn to use the tool
  - They can engage in informed dialogue about their digital-systems uses and requirements

(Starred bullets are considered potential core elements, depending on discipline)

Attendees at the Curricula Development Workshop, as well as at OGF22, will be providing examples of how topics listed in the Stage 2 curriculum would be taught to students from different disciplines.

STAGE 3: Advanced Methods

Stage 3 equips students to choose, configure, parameterise and compose tools in e-Science. As with Stage 2, this stage would be considered optional. Stages 1 and 2 are prerequisites for entrance into Stage 3. This level is domain-specific and involves exposure to a range of tools and to programming.

- **Expect experts in a narrower space as a result of this level**
  - This level is typically very domain specific
- **Can choose, configure, parameterise and compose tools**
- **Able to engage with developers in specifying and evaluating tools**
- **Depending on the subject:**
  - This may include middleware, services & applications
- **Generic tools *may* be part of courses here**
  - E.g. portal / problem solving environments
  - Workflows
  - Grid computing, HTC & (optionally) HPC
  - Concurrency, parallelism & computing architectures
  - Large-scale storage technology
  - High-bandwidth communications

EQF and Learning Outcomes, Level 6 (Bachelors)

In the EU context, it is valuable to keep in mind the European Qualifications Framework (EQF) Learning Outcomes when crafting curriculum. Level 6 corresponds with Bachelors level and involves the following knowledge, skills and competences:

- Advanced knowledge of a field of work or study, involving critical understanding of theories and principles
- Advanced skills, demonstrating mastery and innovation, required to solve complex and unpredictable problems in a specialised field of study
- Competence to manage complex technical or professional activities or projects, taking responsibility for decision-making in unpredictable work or study contexts

Prerequisites and Educational Goals for a Masters Course in e-Science

After completion of the Masters course, students will have a high-level understanding of applications in e-Science and will also have skills in data management, programming and trans-domain communication.

It was agreed that the Masters course should not be linked to the undergraduate levels. Prerequisites would include:

- Mathematics and science competencies (calculus and statistics, numerical, analytical and technical understandings).
- A substantial part of Stage 2 competencies would be required, but not necessarily through taking courses associated with Stage 2.
- Simple programming

Completion of a final project would be a key element of the Masters degree. This project is domain-specific and demonstrates key learning goals.

COURSE CONTENT:

- **Understanding e-Science**
  - collaborative working environments
    - ethics
    - tools
    - interpersonal protocols (communication in remote communication tools)
  - solving larger problems beyond local resources
    - scale of problems
    - broad examples from different disciplines
  - distributed computing for e-Science
    - infrastructures
    - case studies in e-Science
    - things you can do with e-Science, types of problems and how they map to different infrastructures, etc.
    - network comms and implications thereof

- **Data Management**
  - storage
  - movement
  - provenance
  - life-cycle
  - validation
  - security
  - schemas / data formats
  - documentation

- curation

Examples can be domain specific

- **Programming for e-Science**
  - loosely-coupled programming (includes communications, networks issues, workflows...)
  - programming to APIs
  - Code re-use & component publishing, API production
  - code maintenance, versioning, etc.
  - technical documentation for re-use
  - standards
  - programming environments
  - security
  - introduction to existing CS methods & concepts

- **Presentation & Communications skills**
  - Trans-domain communication skills
    - simple guidelines: e.g. don't use jargon or acronyms, etc.
  - case-studies of failures
  - prepare presentation, for someone outside of your domain
  - give various presentations, individual and group presentations.
  - user documentation
  - shared reports, shared documentation, etc.
  - wikis, blogs, etiquette, etc.
  - requirements gathering

- **Final project**
  - must demonstrate key learning goals of the course
  - domain-specific
  - appropriate supervisor who suggests topic
  - individual project
  - tangible product at the end of it
  - assessment via
    - project introduction presentation
    - demonstration of application
    - project report
    - possibly some interim reports, etc.
    - diary/blog of progress
  - possibly produce a research paper from the dissertation
  - trans-domain aspect whereby the student must explain their work so that it can be understood by someone from a different background
  - literature review
  - basic project management

- research methods introduction lecture before they begin the project

Ideally the following should also be incorporated into the project

- showing composition of existing tools as well as writing their own code
- with some collaborative aspect, have to talk to or work with someone if possible
- requirements gathering should be included (if appropriate)

EQF Level 7 (Masters)

EQF Learning Outcomes at Level 7 correspond with Masters courses and can be used as a reference when considering content for a Masters course in e-Science. Knowledge, skills and competencies at this level include:

- Highly specialised knowledge, some of which is at the forefront of knowledge in a field of study, as the basis for original thinking and/or research.
- Specialised problem-solving skills required in research and/or innovation in order to develop new knowledge and procedures and to integrate knowledge from different fields
- Competence to manage and transform work or study contexts that are complex, unpredictable and require new strategic approaches

Existing Masters Courses in Grid Computing and e-Science: comparing content

The ICEAGE website lists Masters courses on offer at universities worldwide (the list is not yet a comprehensive record), including a general description of their content. The University of Edinburgh MSc in e-Science curriculum content is provided here as an example to compare with content discussed at the workshop:

University of Edinburgh MSc in e-Science:

Semester 1 – Distributed Computing for e-Science 1, Software Engineering with Objects and Components, Introduction to Scientific Data, Programming for e-Science.

Semester 2 – Distributed Computing for e-Science 2, Software Architecture, Process and Management, Topics in e-Science, Project Preparation (e-Science) + four optional courses (for example, in Informatics, Physics, GIS)

Concluding Comments

The Curricula Development Workshop was a valuable first step in clarifying content for distributed-systems thinking and e-Science courses at both undergraduate and graduate levels, but much more work needs to be done. The workshop has successfully set forth a framework which can be discussed and developed by educators. In order to progress work begun in Brussels, a further workshop was tentatively proposed during OGF22, to be held just prior to OGF23 in Barcelona (June 2008). Continued elaboration of this curricula framework is vital to the international development of e-Science education.

Appendix B – Grid Education Curricula and Textbook Development Resources

1) Grid Technology Cookbook, SURA
   http://www.sura.org/cookbook/gtcb

2) GridForce Project
   Bina Ramamurthy, SUNY at Buffalo
   http://www.cse.buffalo.edu/faculty/bina/gridforce/first.htm
   http://www.cse.buffalo.edu/faculty/bina/

3) International Workshop on Collaborative and Learning Applications of Grid Technology and Grid Education, 2005 and 2006.
   http://gsic.tel.uva.es/clag/clag2006.html

4) ACM Curricula Recommendations: http://www.acm.org/education/curricula.html

   SIGCSE, ACM Technical Symposiums on Computer Science Education, 2005-2007 (also upcoming 2008 Symposium, "Diversity through accessibility", 12-15 March, Portland OR)
   http://portal.acm.org/browse_dl.cfm?linked=1&part=series&idx=SERIES307&coll=portal&dl=ACM&CFID=21520226&CFTOKEN=81262262

5) IEEE Computer Society Computing Curricula Series
   http://www.computer.org/portal/site/ieeecs/menuitem.c5efb9b8ade9096b8a9ca0108bcd45f3/index.jsp?&pName=ieeecs_level1&path=ieeecs/education/cc2001&file=index.xml&xsl=generic.xsl&

6) BCS Education and Training Forum
   http://www.bcs.org/server.php?show=nav.6042

Endnotes

---

[1] See OGF wiki at http://forge.gridforum.org/sf/wiki/ for the full list of documents

[2] National Cyberinfrastructure Council, *Cyberinfrastructure Vision for 21st Century Discovery*, NSF 07-28, March 2007 (URL http://www.nsf.gov/pubs/2007/nsf0728/index.jsp/

[3] See Computing Business News, "Review 2007: IT skills and careers", Glick, Bryan, Computing 20 December 2007, at http://www.computing.co.uk/computing/news/2206086/review-2007-skills-careers

[4] See Contractor UK News, "Europe faces 'imminent' IT skills crisis", 22 July, 2005 at http://www.contractoruk.com/news/002206.html

[5] ICT Skills Monitoring Group, Synthesis Report, "e-Business and ICT Skills in Europe", 2002 at http://ec.europa.eu/enterprise/ict/archives/index_en.htm and see also the EC Enterprise and Industry webpage at http://ec.europa.eu/enterprise/ict/policy/ict-skills.htm

[6] European e-Skills Forum, Synthesis Report, "e-Skills for Europe: towards 2010 and beyond", pp. 5-6, September 2004 at http://ec.europa.eu/enterprise/ict/policy/doc/e-skills-forum-2004-09-fsr.pdf

[7] See ComputerWorld, "The 8 Hottest Skills for '08", Hoffman, Thomas, 31 December 2007 at http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=308800

[8] See IT News, "Skills shortage threatens Australia's ICT growth", 25 February, 2008, at http://www.itnews.com.au/News/70796,skills-shortages-threatens-australias-ict-growth.aspx

[9] Thibodeau, Patrick, "The goods for grid", TechWorld, 18 October, 2004 at http://www.techworld.com/opsys/features/index.cfm?featureid=924

[10] Phil Wadler, personal communication to Malcolm Atkinson, March 2008

[11] See "Education at a Glance 2007: OECD Indicators", OECD, 2007 at http://news.bbc.co.uk/1/shared/bsp/hi/pdfs/18_09_07_oecd.pdf and see "Constructing Knowledge Societies: New Challenges for Tertiary Education", A World Bank Report, 2002 http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,contentMDK:20283509~menuPK:617592~pagePK:148956~piPK:216618~theSitePK:282386,00.html

[12] i2010 in Context: ICT and Lisbon Strategy, Europe's Information Society, EC Commission, at http://ec.europa.eu/information_society/eeurope/i2010/ict_and_lisbon/index_en.htm

[13] See EC COM (2006) 816 final, Implementing the Renewed Lisbon Strategy of Growth and Jobs, "A Year of Delivery" at http://ec.europa.eu/growthandjobs/pdf/1206-annual-report-en.pdf and EC COM (2005) 118 final, Building the ERA of knowledge for growth, p. 3 at http://ec.europa.eu/transparency/regdoc/liste.cfm?&type=1&annee=2005&numero=118&ElementsPerPage=20&tri=cote&CL=en

[14] The Glossary of Terms will soon be entered into the editorial process for publication as an OGF information document. Currently it can be accessed in its draft form on the OGF wiki at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/Definitions

[15] Definition of e-Science taken from: Atkinson *et al.*, "Century-of-Information Research – a Strategy for Research and Innovation in the Century of Information (CIR3), January 2008, http://wikis.nesc.ac.uk/escienvoy/Century_of_Information_Research_Strategy_%28CIR%29:_a_strategy_to_meet_the_research_challenges_and_opportunities_in_the_century_of_information

[16] ACM Curricula Recommendations at http://www.acm.org/education/curricula.html

[17] See Jeannette M. Wing, "Computational Thinking" in Communications of the ACM, Viewpoint, March 2006/Vol. 49, No. 3, pp. 33-35.

[18] See INSPIRE (Infrastructure for Spatial Information in Europe) homepage at http://www.ec-gis.org/inspire/

[19] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html

[20] Contributors to the 2nd ICEAGE Forum concluded that, "the response should be a …systematic development of a professional discipline, with a body of knowledge and professional practices that will lead to reliable, cost-effective and predictable distributed systems projects and operations", in ICEAGE D1.F2: Second Forum Report, 22/5/07. See also EU Directive 89/48/EEC, which addresses recognition of professional qualifications across EU Member States, at http://ec.europa.eu/internal_market/qualifications/general-system_en.htm

[21] See OGF wiki, "Towards Professional Grid Certification", draft document at https://forge.gridforum.org/sf/go/doc14419?nav=1

---

[22] See SURA Grid Technology Cookbook at http://www.sura.org/cookbook/gtcb/

[23] See European Commission Information Society Activities at a Glance for information on Europe's Information Society Digital Infrastructures programmes at http://ec.europa.eu/information_society/activities/index_en.htm

[24] Sinnott, Stell and Watt refer to the "fluidity of the technological landscape" so that "grid technology and associated standards are perpetually evolving with new recommendations and software from standards bodies and solution providers", in Sinnott, R.O., A.J. Stell and J.P. Watt, "Experiences in Teaching Grid Computing to Advanced Level Students", National e-Science Centre, University of Glasgow at http://eprints.gla.ac.uk/3555/01/sinnott3555.pdf

[25] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html

[26] For information about JSDL please see http://jsdl.sourceforge.net/index.html, for GSM please see http://www.gsmworld.com/about/index.shtml, for GridFTP see http://www.globus.org/grid_software/data/gridftp.php and for SAGA please see http://sourceforge.net/projects/saga-gis/

[27] Sinnott, Stell and Watt highlight the problem, explaining that "understanding the technical and non-technical aspects associated with security is crucial, not least due to the degree of trust between resource providers and the potentially highly distributed remote end users", in Sinnott, R.O., A.J. Stell and J.P. Watt, "Advanced Security Infrastructures for Grid Education", National e-Science Centre, University of Glasgow at http://www.nesc.ac.uk/papers/staff/AdvGridSec-OrlandoFinal.pdf

[28] Please see JANET Roaming Policy, v. 1.47 – 11 April 2006, http://www.ja.net/services/network-services/roaming/documents/policy.pdf (JRS or JANET Roaming Service, is the UK branch of the eduroam confederation), JANET Roaming home page, http://www.ja.net/services/network-services/roaming/, University of Bristol technical specification compliance case study document on JANET Roaming documentation page (etc.) as well as the university's acceptable use policies (which includes university-specific regulations for use of computing facilities, code of conduct and JRS usage policy) at http://www.wireless.bris.ac.uk/wordpress/?page_id=20, and European eduroam confederation policy, Version 1.1, GEANT2 JRA5 deliverable "Roaming Policy and Legal Framework Document – Part 2", http://www.geant2.net/upload/pdf/GN2-06-080v4-Deliverable_DJ5-3_2_Roaming_Policy_and_Legal_Framework-Part2_20060719163405.pdf

[29] Please see OGF wiki, "Experiences and Issues with t-Infrastructure" draft document at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/TInfrastructureExperiences

[30] Oxford University has delved into the issue of IPR in grid computing environments through the IMaGE Project, which particularly focuses on the complexities of sharing medical data. The project has examined eDiaMoND, the UK eScience Digital Mammography National Database, which is being developed through grid technology applications. The IMaGE analysis teases out issues that could be relevant when considering IPR models to suitably frame sharing in EU grid computing and grid education. See D'Agostino et. al., "On the Importance of Intellectual Property Rights for eScience and the Integrated Health Record", Oxford Projects, IMaGE and http://www.oerc.ox.ac.uk/activities/projects/index.xml?ID=image

[31] See Foundation for Information Policy Research, Text of Directive 2001/29/EC at http://www.fipr.org/copyright/eucd.html

[32] See the Berne Convention text at http://www.law.cornell.edu/treaties/berne/overview.html and TRIPs page at http://www.wto.org/english/thewto_e/whatis_e/tif_e/agrm7_e.htm

[33] Please see OGF wiki, "IPR for Grid Education and Training" draft document at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/IPRForGridEducationTraining

[34] ICEAGE: www.iceage-eu.org/library and EGEE: http://egee.lib.ed.ac.uk

[35] See OGF wiki, "Towards Professional Grid Certification" draft document at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/TowardsProfessionalGridCertification

[36] Please see NorduGrid NGIn at http://www.nordugrid.org/ngin/ and Nordic eScience Strategy Document at http:// www.**cs.umu.se**/~elmroth/papers/**nordic_escience**_final.pdf

[36] Please see ICEAGE Summer School webpage at http://www.iceage-eu.org/v2/affiliated%20summer%20schools.cfm and current ICEAGE International Summer School on Grid Computing 2008 at http://www.iceage-eu.org/issgc08/index.cfm

---

See also ICEAGE International Winter School on Grid Computing at http://www.iceage-eu.org/iwsgc08/index.cfm

[37] See OGF wiki, Existing Courses at https://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/ExistingCourses

[38] See EGI website for current information on NGI development in each Member State:

[40] See EGI Knowledge Base Main Page at http://knowledge.eu-egi.eu/index.php/Main_Page

[39] See ESFRI Roadmap at http://cordis.europa.eu/esfri/roadmap.htm

[40] See ICEAGE Forum Agenda Geneva, Third Forum Meeting, EGEE Conference at http://www.iceage-eu.org/events/forumMeetingGeneva.html

[43] See ChinaGrid at http://chinagrid.hust.edu.cn/rms/grid_introduce/introduce_detail.php, Naregi Japanese Research Grid Project at http://www.naregi.org/index_e.html, RedIris at http://www.rediris.es/index.en.html, and KAREN at http://www.karen.net.nz/home/

[42] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html

[43] See OGF wiki, “Towards Professional Grid Certification”, draft document at https://forge.gridforum.org/sf/go/doc14419?nav=1