EDUCATION AND TRAINING TASK FORCE
Report

9 June, 2008*

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 e-IRG ETTF can play a significant role to foster work that will contribute to growth in the area of European e-Infrastructure education and training.

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. There are already demonstrations showing the benefits to engineering, medicine and the creative industries as well. 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. e-Infrastructure provides an environment for coping with these large data sets and for sharing data across regions. An investment in educating people in this technology, then, is an investment that will strengthen our economies and societies. In order to deliver e-Infrastructure education and training successfully in the EU, we must develop a policy framework that will ensure shared responsibility and equivalent training in the field. This document focuses primarily on the current state of grid and e-Science education, introducing key challenges and the opportunities available to educational planners that serve as a starting point for further work. It then proposes strategies and policies to provide a supportive framework for e-Infrastructure education and training.

The ETTF Report concludes with policy recommendations to be taken forward by the e-IRG. These recommendations address issues such as the level of Member State investment in e-Infrastructure education, the harmonisation of education in distributed-computation thinking and in the use of e-Infrastructure and the development of standards for student and teacher identification, for the sharing of t-Infrastructure (and training material) and for accreditation.

Acknowledgements

The development of the ETTF Report was supported by the ICEAGE EU FP6 Project (number RI-026637) and EGEE II.

* Revised in accord with comments at Lugano meeting 26 June 2008, which accepted the report subject to these minor rewordings.

Long Report Version 6, final, 09/06/2008
Malcolm Atkinson, Elizabeth Vander Meer, David Fergusson and Morgane Artacho

CONTENTS

Report Purpose 1
Abstract 1
Acknowledgements 1
1. Executive Summary..... 3
2. Motivations for Investing in e-Infrastructure Education and Training: Foundation of a Knowledge-Based Economy 6
    2.1 The skills and knowledge shortage 6
    2.2 Optimisation of the use of e-Infrastructures..... 7
    2.3 Benefits for industry and academia..... 8
    2.4 EU policy provisions - education, research and ICT 8
3. Challenges and Requirements for e-Infrastructure Education and Training 11
    3.1 Curricula and textbook development 11
    3.2 General Expertise in e-Science: grid computing and computer science 16
    3.3 Teaching and the fluidity of the technological landscape 17
    3.4 Disparate educational policies: harmonisation and security 17
    3.5 Sharing training infrastructure 18
    3.6 IPR and sharing..... 19
    3.7 Training-specific challenges and requirements..... 19
    3.8 Impact of standards on education and training..... 20
4. Opportunities and Existing Structures for Education and Training..... 20
    4.1 Existing educational machinery: curricula, t-Infrastructure and security 21
    4.2 NGIs and the EGI – providing infrastructure for education and training..... 21
    4.3 Embedded e-Infrastructure in national educational operations, plans and policies 22
5. A Framework for Policies Required: Models to Consider 22
6. Suggested Strategies for Development and Provision of Education and Training..... 23
    6.1 Curricula development..... 23
    6.2 Certification 24
    6.3 Promote the sharing of resources 24
    6.4 Develop relationships..... 24
7. Policy Suggestions and Conclusion..... 24
8. Contributors..... 26
9. Appendices 28
    Appendix A - Definitions..... 28
    Appendix B – ICEAGE Curricula Development Workshop Report..... 29
    Appendix C – Grid Education Curricula and Textbook Development Resources..... 39

1. Executive Summary

The e-IRG Education and Training Task Force (ETTF) was established to address concerns surrounding the current state of e-Infrastructure education and training in the EU. The United States National Science Foundation (NSF) presents a definition of 'cyberinfrastructure' which can be used to describe the various components of e-Infrastructure so that the words are interchangeable[1]:

> 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 ETTF. EU investments in e-Infrastructure require adequate investments in e-Infrastructure education in order to allow EU Member States to fully develop and exploit these technologies for academic, industrial, governmental and medical research and innovation. Providing coordinated education across Member States will, therefore, lead to a strengthened European Research Area (ERA) and thus to the EU's ability to compete in the knowledge-based global economy as seamless e-Infrastructures are developed by newly educated professionals highly competent in the field.

As a means of confronting and correcting the skills and knowledge shortage apparent in this area of computing technology[2], the ETTF Report argues for further investment in education and training; it envisions the embedding of education and training into normal academic training in Europe. This Report presents a list of motivations that justify this vision. 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 growth of e-Infrastructure education across the EU.

Motivations for increased investments in e-Infrastructure education and training

Coordinated development of e-Infrastructure across the EU is vital to maintain Europe's competitive edge in the knowledge-based economy, supporting advances in science, industry and education. Four key motivations associated with this point and providing the impetus for increased investments in education and training have been identified by the ETTF and presented in the report:

1) a skills and knowledge shortage
2) optimisation of the use of e-Infrastructure
3) benefits to industry and academia
4) the relationship of e-Infrastructure education and development to EU policy provisions

Challenges for education

The ETTF vision of coordinated, embedded EU e-Infrastructure education can only be realised after understanding the current state of education in this field. Challenges that we face due to limitations in existing methods and tools, particularly identified in grid computing education, include poorly developed curricula and textbooks to support that curricula, lack of students motivated and prepared for e-Science and e-Infrastructure courses, the fluidity of the technological landscape, lack of general expertise, disparate educational policies across universities and countries, which create security and access problems, and lack of a shared t-Infrastructure or solid IPR framework. The ETTF Report explores these problems in depth and suggests ways to confront them.

Opportunities for education

We can also identify opportunities in existing grid education, those methods and tools that can be used and developed in order to realise the ETTF vision. Undergraduate courses, Masters courses, Doctoral training colleges and summer schools are currently run in various Member States throughout the EU, providing working examples of teaching modes and curricula (best practice).[3] National Grid Initiatives (NGIs) are being developed in 37 European countries, providing national infrastructures for education. The European Grid Initiative (EGI) has the potential to support coordination of NGIs so that national education and infrastructures are harmonised into *EU* education and infrastructures. And finally, a number of EU Member States have already successfully embedded e-Infrastructure education into their national policies.[4] The ETTF proposes ways to build on these existing methods and frameworks in order to achieve the goal of coordinated, embedded e-Infrastructure education across the EU.

Strategies and policies

Strategies presented in the ETTF Report follow on from the challenges and opportunities already identified. As noted, we currently face challenges in grid education in the area of curricula development, so this must be a priority when considering strategies that will embed and to some extent “standardise” e-Infrastructure education throughout the EU. The list of strategies also includes means to develop certification, resource sharing and relationships (for instance, between NGIs and the EGI) in order to achieve the ETTF’s goal.

If we consider what strategies are required to embed e-Infrastructure education in normal academic training in Europe, we can then create a policy framework to support and advance the ETTF vision. The Report identifies the need for two kinds of policy, one directed at providers of education and one for teachers and students. Recommendations suggesting policies that the e-IRG should develop conclude the ETTF report.

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 Europe. 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, the EU will flounder in its attempts to become leader in the knowledge-based economy. To prevent this from happening, the ETTF Report provides concrete suggestions for strategies and policies which will extend and advance e-Infrastructure education across Member States and thus allow the EU to fully exploit the technologies at its “fingertips”.

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 the ERA’s capacity in this field.

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

2. Motivations for Investing in e-Infrastructure Education and Training: Foundation of a Knowledge-Based Economy

OECD and World Bank country studies have confirmed an obvious correlation between investment in education and quality of life and GDP.[5] There are economic benefits to educating EU 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 information and communications technologies (ICT) infrastructures now exist in most universities.[6] The EC Benchmarking Reports identify the spread of broadband and internet in homes and businesses as well as ICT use in almost one hundred per cent of European schools. ICT is recognised by the EC as key to a knowledge-based economy and social cohesion, and so it must have a place in education and training.[7] 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. We can see that further investments in e-Infrastructure education and training would be beneficial within this social context. The most pressing motivations for increased investments are listed below.

2.1 The skills and knowledge shortage

There is currently a skills and knowledge shortage in the e-Infrastructure industry (including grid computing) which has led to *a definite crisis* resulting in distributed system failures.[8] We need more individuals educated in e-Infrastructure technology in order to address the crisis. We also need an increase in the skilled workforce to fully exploit the opportunities brought by e-Infrastructure and e-Science.

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.[9] Prior to this, the Leitch Report had issued a more general statement at the end of 2006, urging the UK to devote significant energies to strengthening adult skills in economically valuable areas[10]; skills in the use of e-Infrastructure technologies would certainly fall under this category.

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.”[11] . The European Commission ICT 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.[12] The European e-Skills Forum, established by the European Commission, reported on the impacts and reasons for the current situation[13]:

> 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.*

Concerned computer scientists have particularly pinpointed a lack of expertise in grid computing[14], explaining that:

> 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.
.

2.2 Optimisation of the use of e-Infrastructures

e-Infrastructure technologies such as grid computing involve the potential risk of poor return on investment. A compelling example, which applies to any research infrastructure (ESFRI) is that it takes years of training to get the best out of facilities.[15] Gaining the best from e-Infrastructure is not simply running the most jobs or the largest volume of data. Nor is it about engaging the 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 EU does not want to fall behind countries such as the U.S., who already invest significantly in research and technology development. In the EC communication setting the foundation for creation of the ERA (2000), statistics revealed that Europe's expenditures on research and development in 2000 were significantly less than those of the U.S. and Japan. The EU was also out-competed when considering the number of researchers employed by Member States.[16] The Seventh Framework Programme (FP7) is now in place to provide funding for innovative ICT technologies such as grid computing, to allow Europe the chance to surge ahead of its competitors in both research and industry. This point regarding level of investment in R&D needs to be stressed in discussions with funding bodies, and examples of e-Infrastructures technologies in action should be marshalled to show its potential and thus the importance of increased investment in e-Infrastructure education.

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 transform the ERA 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.

2.3 Benefits for industry and academia

Both industry and academia benefit from e-Infrastructure, or grid computing, outputs. For instance, in finance and online industries, use of e-Infrastructure has already become integral to businesses' functioning and thus ultimately to 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. e-Infrastructure education prepares students for employment in these fields. In academia, research is becoming increasingly more collaborative and use of e-Infrastructure facilitates information and resource sharing and collaborative action.[17] Today's research, much like industry, also tends to generate vast amounts of data that needs to be properly managed (collated and analysed).

Phil Wadler, Professor of Theoretical Computer Science at University of Edinburgh[18], observes that:

> 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).

Industry and Academia: Elements of the Knowledge Triangle

Use of e-Infrastructure assists in simulation, analysis and better management of data.[19] e-Infrastructure technologies have the potential to foster stronger links in the knowledge triangle, which includes research, education and innovation; in order for the EU to become leader in the knowledge-based economy, Europe "must become better at producing knowledge through research, diffusing it through education and applying it through innovation."[20] Use of e-Infrastructure can make this possible, enhancing and expanding potential in industry and academia.

2.4 EU policy provisions - education, research and ICT

EU policies aim to serve the objectives of the Lisbon Strategy. The Lisbon Strategy is the key action plan for continued European growth and development. Its primary aim is to assist European countries in their transition to knowledge-based economies**.** Policies and funding relating to Education, Research and Development and Information and Communications Technology (ICT) all pertain to e-Infrastructure because they support the EU's transition to a knowledge-based economy; they support an argument for expanding and advancing grid education in particular and e-Infrastructures more broadly. EU policy provisions recognise the role of Member State cooperation in education and the importance of ICT to facilitate this transition towards a collaborative standard. EU research policies encourage integration of the knowledge triangle in order to ensure that Europe maintains a competitive edge in the knowledge-based economy. As already highlighted, e-Infrastructure provides a platform for innovation in industry as well as in academic research. So promotion of e-Infrastructure education across Europe fulfills the aims of the Lisbon Strategy as well as associated EU policies on education, research and ICT.[21]

EU Education Policy

The Treaty of the European Union (Article 149) recognises the importance of education and formally encourages Member States and educational establishments to cooperate in the field of education. Relying on the cooperation of its Member States, the EU has established education programmes and processes. The Education and Training 2010 Programme is the main umbrella under which various education and training policies and funding have been generated to achieve the Lisbon goals by 2010. The Bologna Process is of the utmost relevance to the Higher Education sector. Its three-cycle system, of Bachelors, Masters and Doctoral degrees, aims at creating a European Higher Education Area by 2010, thus helping the EU Member States compete with the rest of the world.[22]

EU Research Policy

It is undeniable that the best economies in the world invest heavily in research and development. The EU has therefore ensured the running of a strong funding programme (the current Seventh Framework Programme) to help bridge the gap with other competitive economies.

It is vital that Europe completes the creation of the European Research Area. The latest consultation on the Green Paper on the ERA proves the strong policy commitment which emanates from the EU – it is our responsibility to respond to this trend and the various issues listed in the Green Paper. Funding must be targeted to direct convergence of the ERA with other established EU “areas” (European Higher Education Area, European Innovation Area, etc), to integrate the knowledge triangle, since all are mutually supportive of the same endeavour: to ensure that Europe maintains a competitive edge in knowledge and innovation.[23]

The European Strategy Forum on Research Infrastructures (ESFRI) serves as an instrument for policy-making concerning European research infrastructures and facilitates negotiations about initiatives.[24] An important facet of FP7 is to fund ESFRI projects, which proves the need and relevance of such a strategy. There are a substantial number of research areas and projects that will benefit from e-Infrastructure.

In particular, many of the facilities generate large volumes of digital data that need to be accessed and analysed, many depend on computational modelling for experimental design and planning and a select few depend on computational steering. All of these demand advances in skills and a greater number of suitably skilled researchers.

In support of the creation of European infrastructures to advance research, the e-IRG highlighted the need for funding of e-Infrastructure education within the ERA in the recommendations and decisions paper produced during the Finnish presidency (2006): “The e-IRG recommends that the EC support the launch of an ERA-wide activity to coordinate education and training efforts, with an emphasis on the efficient exploitation of e-Infrastructure by EU citizens."

The EU and ICT

The EU is transitioning into a knowledge-based economy via different policies and funding programmes pertaining to the Lisbon Strategy, and this transition brings with it the problem of managing vast amounts of data, which requires cyberinfrastructures or e-Infrastructures (ICT) to access and manage "global" knowledge.[25] As previously noted, e-Infrastructure provides computing frameworks to manage and allow for the wide sharing of this data. Education about and in the use of e-Infrastructure, then, can have significant value within the current social context. In a speech at the CeBIT Trade Fair in Hannover (March 2008), President of the European Commission, Jose Manuel Durao Barroso, stressed the importance of ICT in achieving Lisbon strategy goals[26]:

> The Information and Communications Technologies, in short ICT, industry is a motor for our economies and for our joint prosperity…The ICT industry has therefore a major role to play in the European economy of the 21st Century…crucial to the success of the EU Lisbon Strategy for growth and jobs.

Barroso then listed areas that should receive special attention, in order to strengthen ICT and thus increase European competitiveness, two of which are particularly relevant to education and training provision: investment in human capital and research. As is evident here, the EC fully sanctions greater investments in e-Infrastructure education.

The EC focus on ICT has sparked research to ascertain the prevalence of citizens' access to this technology within the EU. The resulting Benchmarking Reports present concerns about the digital divide that currently exists in Europe. To begin to close the digital divide, the EU has created the i2010 Policy Framework, which promotes the inclusion of all EU citizens in the developing information society. Inclusion can come through adequate educational opportunities in ICT.

The European Commission has not only identified and taken steps to address the digital divide, but it has also more recently recognised skills and knowledge shortages and gaps in ICT. In September, 2007, the Commission issued a communication that will stand as a guide for a long-term e-skills agenda; it follows the formation of the European e-Skills Forum. The need for such an agenda arises from the following acknowledgement[27]:

> A compelling concern is now the reported decline in supply of good graduates from ICT courses. Growing e-skills shortages and mismatches are expected in future.

The EC communication stresses the importance of confronting this crisis in order to meet increasing demand for these skills, the needs of a fast-changing industry which requires up-to-date skills and knowledge and to allow all EU citizens the opportunity to engage with ICT.[28] The long-term agenda proposed within this document in effect endorses the strengthening of e-Infrastructure education while also presenting a general framework that the ETTF can reference in order to more rapidly achieve its goals.

Training in e-Infrastructures, including grid computing, can be done within existing (ICT permeated) academic structures, which means investment in creating new structures need not be made. EU policy already sanctions and supports this change of existing structures through the programmes and funding initiatives listed above. The question then becomes how to ensure that students are well trained in grid computing, a question which leads into a range of challenges and opportunities for education in this area.

3. Challenges and Requirements for e-Infrastructure Education and Training

This section details the current gaps in grid and e-Science education, as well as in training, that 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 to develop a broader agenda. In order to contribute, please visit the e-IRG ETTF wiki page at:
> https://eirgsp-wiki.grnet.gr/bin/view/Main/TrainingAndEducation

Certain required tools or structures may be missing which could hold back attempts to extend grid and e-Science education across Member States. We have identified the following challenges:

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 needs 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 throughout the EU.

Content Development – Building on the ICEAGE Curricula Development Workshop

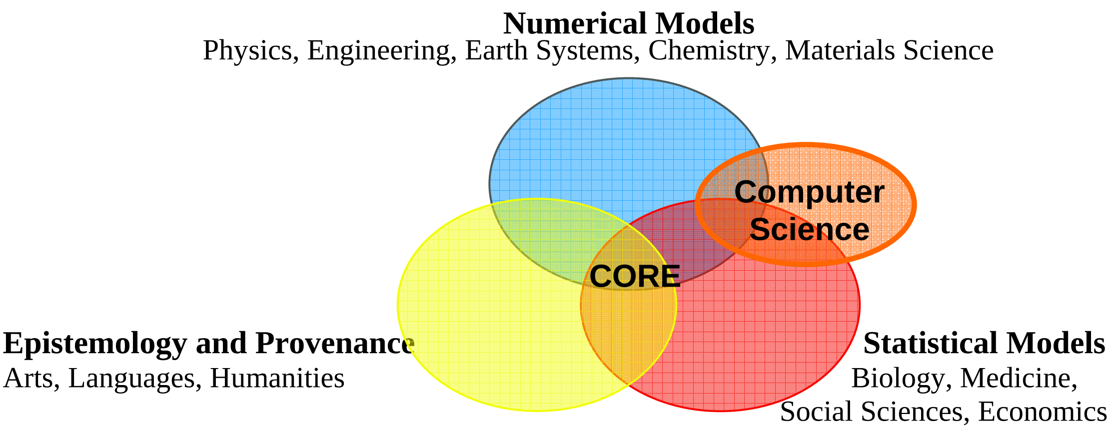
There are existing sources to consider in the area of curricula development. For instance, the ACM produces curricula guides for computer science courses and these can be used as a reference.[29] But the design of curricula specific to grid and e-Science education at present needs considerable focused effort, and this effort has recently been initiated by the ICEAGE Project. 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 e-IRG ETTF (see Appendix B). 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[30]:

> 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 and e-Science education has been developed through OGF-ETTF associated workshops. At OGF22, the ET-CG proposed organising a session at the next OGF devoted to discussion of curricula; this was arranged and took place at OGF23 in Barcelona (June 2008). A further curricula development workshop to follow up on efforts begun in Brussels has been tentatively proposed for Fall 2008, hosted by NeSC at University of Edinburgh.

Proposed Categorisation of Disciplinary Areas



Curricula discussions in various fora have led to the conclusion that education and training in e-Infrastructures should begin with early development of core skills which are taught across disciplines; at undergraduate level, *all* students would learn core digital-information thinking skills. Disciplines applying e-Science would then learn domain-oriented computational thinking skills at undergraduate level, the set of skills dependent upon whether the discipline relied most heavily on numerical models, statistical models or epistemology and provenance. Undergraduate Computer Science students would build on foundational core skills to gain specialised knowledge of distributed systems, data and computational systems and software engineering.

In summary, the proposed categorisation of disciplinary areas would be as follows:

- All disciplines – CORE SKILLS
    - Digital-information thinking
- Disciplines applying e-Science – DOMAIN-ORIENTED COMPUTATIONAL THINKING SKILLS based on:
    - Numerical models
    - Statistical models
    - Epistemology and provenance
- Computer Science – SPECIALISED KNOWLEDGE of:
    - Distributed systems
    - Data systems
    - Computational systems
    - Software engineering

More details of preliminary work on curricula can be found in Appendix B (ICEAGE Curricula Development Workshop Report).

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, Baysian, 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.[31] 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, compose 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 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 observe 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 when 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 by the e-Science courses.

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.

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[32]:

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.

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, 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.[33] Accreditation bodies such as the BCS and UK Engineering Council could play an important role by certifying courses so that students completing these courses have widely recognised degree credentials which are therefore more valuable to potential employers. For the benefit of the ERA, such accreditation needs coordination and mutual recognition across Member States. 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.

It is recommended that a high-profile and broadly representative committee of leading educators across disciplines, ICT experts and leading employers, be formed to expedite the creation of the curricula goals and principal topics. This should be launched and concluded by major European conferences drawing attention to the educational priorities and opportunities in the field.

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 not only in terms of content for future grid computing textbooks originating in the EU, but also in terms of the collaborative efforts involved in its creation.[34]

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 textbooks 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 the EU academic research community and help it compete as a whole with other academic communities such as those established in the United States.

It is necessary to prime and stimulate an incremental growth in the EU's e-Science educational capability. This has already started in some Member States, partly due to the effects of the Information Society Digital Infrastructures programmes[35]. 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 across the ERA 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 grid 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 grid 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.[36] 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 response across the EU, leading to advancements across disciplines and an overall increase in competitiveness across EU Member States. 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.[37] 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 ERA. 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 (all EU citizens).

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[38], 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.[39] 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 e-Infrastructure and grid education and training policies for the EU.[40]

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 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.[41]

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 help. Collecting example data of predicted hurricane tracks is probably relationally 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 adoption 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. It may also be possible to support the required data accesses rates and reliability by just storing a few copies for a European-wide education programme.

3) <u>Experience of interpreting medical images.</u> 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 stimulate 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 across the ERA.[42] 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).[43] Considering a wider context, the Berne Convention is well-established and addresses the issue of copyright, as does TRIPs, within the World Trade Organisation (WTO) agreements.[44] 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 tackled within the Education and Training Community group at the Open Grid Forum. 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 throughout the EU.[45]

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 EU-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 established across Europe, 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 been done by the OGF ET-CG to suggest types of certificates, based on skill sets.[46] 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. 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.

4. Opportunities and Existing Structures for Education and Training

It is important to understand the existing state of grid education in order to know what options are out there for educational planners and how to proceed. We next identify existing tools and infrastructure and what can develop from these.

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. NorduGrid has launched 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.[47] This coordinated effort should be watched and reviewed as a possible example for other European regions to follow.

With some exceptions, there appear to be few coordinated efforts across universities to work together on provisions for the Masters courses. Other (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.[48]

As noted under challenges, 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 primarily 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. They could ensure that courses are certified by accreditation (industry and professional) bodies. In the UK, standards have been set by the Engineering Council, for instance, leading to a professional chartering framework, and the BCS provides a framework through its professional IT examinations.

The EQF Learning Outcomes

The adoption of a European Qualifications Framework (EQF) has been valuable in terms of qualifications recognition across the EU and this will be of assistance when considering harmonisation of e-Infrastructure education (the summer school experience speaks to this). As part of the Bologna Process, the EQF translates all types of qualifications, so that each Member State can identify and recognise equivalencies in qualifications gained under different national systems. The EQF defines eight learning outcome levels and lists knowledge, skills and competences that pertain to each level. This framework allows EU citizens greater mobility and it provides each EU Member State with education and training reference points (links to all other Member States) but also flexibility when crafting curricula and associated qualifications schemes.[47] If we return to the UK qualifications examples, we find that both sets of standards (undergraduate and postgraduate) are recognised by the UK National Qualifications Framework (NQF) and translate into EQF levels 6 and 7.

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 across Member States, so that mobility is facilitated (access across Member State boundaries), resulting in *EU* education in grid computing. Development of an EU-wide infrastructure would advance the sharing of curricula, qualifications and teaching methods.

Existing NGIs and the EGI could provide foundational infrastructure for grid education. There are developing NGIs in 37 European countries which could in principle provide infrastructure for grid education.[49] 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.

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 the 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.[50] European e-Infrastructure integration also has to consider HPC and the 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[51] 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 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.[52] 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) and can serve as references for best practice.[53] We need to identify whether there are models that would allow this to happen elsewhere in EU Member States, and across Member State boundaries.

5. A Framework for Policies Required: Models to Consider

When considering opportunities for educators, we have to keep two models in mind: one based on harmonisation and one that allows grid and e-Science 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 or discipline requirements faster, and flexibility to better respond to a rapidly changing domain. We should consider both models when formulating strategy and policy recommendations. The optimal solution for the ERA will certainly contain elements of coordination and elements of evolutionary competition.

6. Suggested Strategies for Development and Provision of Education and Training

The ETTF Report sets out options to increase engagement with e-Infrastructure technology, and distributed computing in particular, throughout the EU. This document has reviewed the current state of grid and e-Science education, presenting related challenges and opportunities and suggests strategies and policies that can lead to the embedding of education and training into normal academic training in Europe.

6.1 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) Establish a committee/body of leading educators across disciplines to expedite the creation of the curricula goals and principal topics, launched and supported by major European 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 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 offered in each EU Member State will be compiled through your input on the e-IRG ETTF wiki page at https://eirgsp-wiki.grnet.gr/bin/view/Main/TrainingAndEducation. Please add to the list with courses offered in your Member State).

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 C 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, considering EC benchmarking and other report recommendations.

6.2 Certification

Encourage certification of courses by professional accreditation bodies such as the BCS and UK Engineering Council in the UK, for example. Consider the role of the EQF in qualifications recognition in the EU. Establish international recognition of the accreditation in each Member State.

6.3 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.[54]

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

6.4 Develop relationships

Look at national and international e-Infrastructure to support education and work with the EGI to promote NGI provision and coordination through establishment of an international network. Consider the potential role of the European Institute of Technology (EIT)[55], multinational facilities and the ESFRI roadmap.

7. Policy Suggestions and Conclusion

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

a) Policy for providers of education—common rules to address issues arising from the sharing of ideas, software and computing.

b) Policy for users/students—common rules to address issues arising from **equipment use** (t-Infrastructure and the bureaucracy surrounding its use, to avoid student queues) including conditions of use (so that systems do not crash) and **mobility and the need for access** (to allow student mobility and continuity of work, for instance, with PhD students).

The ETTF therefore recommend that e-IRG develop policies on the following issues:

1) Recommendations as to the level of investment necessary in Member States in order to provide education in the use of e-Infrastructure.

   Suggestion: The national, regional or European investment in relevant education and training, which is primarily oriented to equipping graduates to use e-Infrastructure well, should be comparable with the investment that is going into e-Infrastructure provision. The recommended strategy for achieving this is to persuade the universities to adapt their curricula in order to prepare their graduates. The significant investment is justified based on the crisis we currently face. Unless there are adequate numbers of people schooled in the creation, use and further development of e-Infrastructure technologies, Europe and its Member States 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 the outflow of skilled individuals inevitably involves commitment in the form of funding. That funding should be catalytic in order to encourage the changes in university curricula.

2) Recommendations to align the development of distributed-computation knowledge and skills.

   Suggestion: Academic institutions, particularly universities, should build on proposals for undergraduate and postgraduate courses in digital systems judgement and e-Science formulated at curricula development workshops. Further work on developing and agreeing those curricula should be undertaken by individuals proposing to teach this material and by relevant professional bodies. The goal of this alignment should be mutual recognition and understanding, not uniformity.

3) Recommendations as to the harmonisation of education in the use of e-Infrastructure.

   Suggestion: Professional bodies, e.g. the Royal Society of Chemists and the Institute for Engineering and Technology in the UK, should identify target attainments in the exploitation of e-Infrastructure for their profession and should harmonise these across the ERA in accord with the Bologna framework. The goal of this harmonisation is not uniformity of skills and knowledge. Rather, it is a common framework to support student mobility and mutual recognition of qualifications, particularly where they are used to appoint staff to positions in the use or operation of critical e-Infrastructure.

4) Propose standards for student and teacher identification that would enable access to educational grid facilities and authorization/management of the resources used.

   Suggestion: A task force, set up by the e-IRG, EGI and GÉANT should build on the eduroam protocols to extend them to cover student use of collaboration facilities and multi-site t-Infrastructure.

5) Propose standards for sharing training material and t-Infrastructure between institutions.

   Suggestion: Those in the EU developing e-Infrastructure courses should build on creative commons for sharing all educational material and the EGI should mediate agreement between NGIs on sharing t-Infrastructure.

6) Establish a system for agreeing standards that accredit workers who design, build, operate and support e-Infrastructure so that qualifications are recognised across the ERA.

   Suggestion: The bodies in Member States that accredit technical skills and knowledge should adapt the proposals developed by the OGF ET-CG working group[56] in order to develop European-wide recognised qualifications.

This initial report for the e-IRG shows that there is a considerable need for increased education and training in the exploitation of e-Infrastructures. There are also skill shortages to be met in order to provide and operate those e-Infrastructures on a scale that will be required.

The report recommends increased investment in education and training and inter-state collaboration on its provision, organisation and evaluation.

There remains a great deal of work to be done in order to define the requirements more precisely, to develop internationally recognised curricula, to support student and teacher mobility and to deliver courses on a sufficient scale to meet European needs.

8. Contributors

Authors

Malcolm Atkinson, Director, e-Science Institute and e-Science Envoy, Edinburgh
mpa@nesc.ac.uk

Elizabeth Vander Meer, Policy Support Officer, NeSC, Edinburgh
evmeer@nesc.ac.uk

David Fergusson, Deputy Director of Training, Outreach and Education, NeSC, Edinburgh
dfmac@nesc.ac.uk

Morgane Artacho, Project Officer ICEAGE, NeSC, Edinburgh

Contributors

The following individuals have made valuable contributions to the content of the ETTF Report:

Torsten Antoni, Giedrius Vilunas, Alberto Di Meglio, Wolfgang Gentzsch, Annu Jylha-Pyykonen, Leo Budin, Fotis Karayannis, Rosa Badia, Lajos Balint, Roberto Barbera, Gaspar Barreira, Dieter Kranzlmueller, Erwin Laure, Marc Ozonne, Peter Kacsuk, Ilmars Slaidins, Peter Sloot, Laimutis Telksnys, Nikos Vogiatzis, Antonio Pina, Dany Vandromme, Giorgio Occhioni, Kyriakos Baxevanidis, Oscar Corcho, Luis M. Vilches Blazquez (contributor to the Curricula Development Workshop Report) and Graham Pryor (also contributor to Workshop Report).

Attendees at the Brussels ICEAGE (OGF-ETTF) Curricula Development Workshop, held at Scotland Europa in February, 2008, were key contributors to the report presented in Appendix B: Amy Apon (University of Arkansas), Mark Baker (University of Reading), Kenny Baird (NCeSS), Kathryn Cassidy (Trinity College, Dublin), Ben Clifford (OSF, University of Chicago), Joy Davidson (HATII, University of Glasgow), Fotis Georgatos (GRNET), Petar Jandric (NeSC, University of Edinburgh) and Fernando Silva (Universidade do Porto).

9. Appendices

Appendix A - Definitions

- **Education –** a long-term institutionalised process using conceptual models and resulting in development of a culture.
- **Eduroam –** Education Roaming, or eduroam, is an infrastructure that allows staff and students to access wireless networks at cooperating universities across the EU (and elsewhere) using their home institution username and password (so they do not have to set up new accounts at institutions they visit). Access is secure and mobile.
- **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.[57]
- **European Grid Initiative (EGI) –** a European level infrastructure based on NGIs, currently in its design phase. The EGI will coordinate NGI interaction and integration in order to improve access to resources and research across the EU. Once developed, it would be considered a key element of the ERA.
- **"grid in a box" –** a local grid infrastructure with easy installation and portability that can be used in universities and other institutions that do not already have developed (or appropriate) grid computing infrastructures.
- **National Grid Initiatives (NGIs) –** national initiatives to connect and expand grid infrastructures in Member States in order to integrate resources for research and allow for coordination, coherence and interoperation for users of grid computing and its applications across disciplines.
- **t-Infrastructure –** e-Infrastructure adapted to the needs of education, trainers and students. Shared t-Infrastructure would be usable by students and teachers throughout the European Research Area (ERA), providing easy access to educational exercises running on (good emulations of) e-Infrastructure.
- **Training –** a short-term process to develop specific skills in a certain technical area.

Appendix B – ICEAGE Curricula Development Workshop Report

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

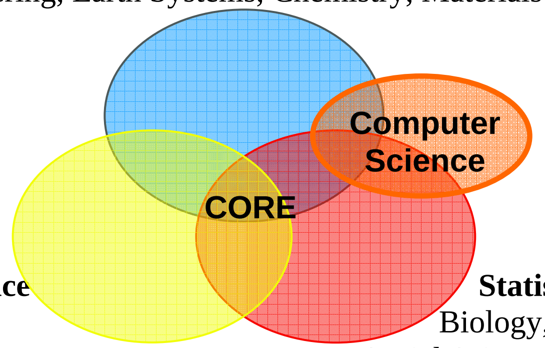
The ICEAGE Curricula Development Workshop, held at Scotland Europa in Brussels from 14-15 February 2008, was successful in proposing an initial framework for undergraduate and Masters level courses in digital-systems thinking and e-Science, which can provide the basis for further work. The workshop was 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).

Numerical Models
Physics, Engineering, Earth Systems, Chemistry, Materials Science



Epistemology and Provenance
Arts, Languages, Humanities

Statistical Models
Biology, Medicine,
Social Sciences, Economics

Workshop attendees focused on developing a framework for e-Science education across disciplines rather than grid education, which would target students within computer science. At the start of the meeting, it was agreed that computer science curricula development would *not* be the focus of discussion. But the group noted that an appropriate computer science curriculum should be developed, in a different forum.

It is understood 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**

    1. 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)
    2. Develop team working using the above
    3. Use digital-communication tools (work with libraries)
        - To coordinate & develop a deliverable
        - Responsibility, legal, ethical & social issues

- Security and safety

4. Critical thinking
5. Use of subject-specific digital resources
    - Scientific data and document data
    - Metadata and controlled vocabularies
    - Proper citation and legitimate use
    - IPR
6. Collaborative behaviour
    - Plagiarism and its detection
    - Drawing on strengths and 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 e-Science 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**. This means that all topics appearing here will not be taught within a Stage 2 Basic e-Science Methods course; for many disciplines, subjects included on the list are already part of their course requirements and could be counted as prerequisites.

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. The order of topics in the list does not denote importance. Asterisks next to topics highlight core elements (but again, this depends on the discipline).

<u>Curriculum Components</u>

Critical thinking (2)*

- Data curation and management*
    - o Data lifecycle, Data Bases, Data models, Metadata, Mark up languages
    - o Controlled vocabularies, Ontologies
    - o Compression and data views

- Using multiple data resources
  - Semantic and provenance

- Process thinking

- Statistical thinking
  - Sampling and error
  - Uses and abuses of statistics

- 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

- Complexity
  - Origins of real-world & system complexity
  - Handling complexity

Responsibility, legal, ethical & social issues (2)*

Presentation and interpretation of data

- Visualisation
- Interdisciplinary professional communication*

Image analysis

- Derivation of information

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

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.

1. **Expect experts in a narrower space as a result of this level**: this level is typically very domain specific
2. **Can choose, configure, parameterise and compose tools**
3. **Able to engage with developers in specifying and evaluating tools**
4. **Depending on the subject:** this may include middleware, services & applications
5. **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 (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 communication and implications there-of

- **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). Instead of holding a workshop, a session of the ET-CG at OGF23 was devoted to curricula development and a workshop was suggested for later in the year (Fall 2008) hosted by NeSC in Edinburgh. Continued elaboration of this curricula framework is vital to the international development of e-Science education.

Appendix C – 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] National Cyberinfrastructure Council, *Cyberinfrastructure Vision for 21st Century Discovery*, NSF 07-28, March 2007 (URL http://www.nsf.gov/pubs/2007/nsf0728/index.jsp/
[2] For details on this shortage, please see Section 2.1 which provides substantial references.
[3] Please see the ICEAGE list of Masters courses at http://www.iceage-eu.org/v2/msc%20courses.cfm
[4] Details on embedded e-Infrastructures will be provided in Section 4.3
[5] 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
[6] *Studies in the Context of the E-Learning Initiative: Virtual Models of European Universities (Lot 1)*, Final Report to the Commission, DG Education and Culture, PLS RAMBOLL Management, February, 2004.
[7] See *Information Society Benchmarking Report 2005, Benchmarking Access and Use of ICT in European Schools 2006* and EC, Europe's Information Society i2010, Digital Divide http://ec.europa.eu/information_society/eeurope/i2010/benchmarking/index_en.htm http://ec.europa.eu/information_society/eeurope/i2010/digital_divide/index_en.htm
[8] Please see McGettrick et. al., "Special Session, The Current Crisis in Computing: what are the real issues?", SIGCE'06, March 7-10, 2007, Covington, Kentucky, ACM. The authors define the crisis not only in the United States, but worldwide.
[9] 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.
[10] Leitch Review of Skills; Final Report, December 2006 at http://hm-treasury.gov.uk/leitch.
[11] See Contractor UK News, "Europe faces 'imminent' IT skills crisis", 22 July, 2005 at http://www.contractoruk.com/news/002206.html.
[12] 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
[13] 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.
[14] Thibodeau, Patrick, "The goods for grid", ComputerWorld, 18 October, 2004 at http://wwwtechworld.com/features/index.cfm?featureID=924&printerfriendly=1.
[15] See ESFRI Roadmap at http://cordis.europa.eu/esfri/roadmap.htm
[16] EC COM (2000) 6 final, Towards a European Research Area at http://europa.eu/scadplus/leg/en/lvb/i23010.htm .
[17] For example, the myGRID project for life sciences researchers has allowed collaboration and subsequent breakthroughs in the study of Grave's Disease and Williams-Beuren Syndrome. See http://www.mygrid.org.uk/index.php?module=pagemaster&PAGE_user_op=view_page&PAGE_id=62&MMN_position=74:74
[18] Phil Wadler, personal communication to Malcolm Atkinson, March 2008
[19] The e-IRG Roadmap highlights the changing nature of science and research and the role and benefits of e-Infrastructure in supporting new methods in science as well as industry, see pp. 8-9, Introduction, and pp. 49-52, *e-Infrastructures Roadmap*, 2005 at http://www.e-irg.org/publ/#Roadmap.
[20] See EC COM (2006) 816 final, Implementing the Renewed Lisbon Strategy of Growth and Jobs, "A Year of Delivery" and EC COM (2005) 118 final, Building the ERA of knowledge for growth, p. 3.
[21] Integration of resources and collaborations within the EU through use of e-Infrastructure are promoted in the e-IRG Roadmap, as a means to strengthen the European knowledge-based economy, see pp. 12-16, pp. 49-52, *e-Infrastructures Roadmap*, 2005 at http://www.e-irg.org/publ/#Roadmap.
[22] Please see European Commission, Education and Training, The Bologna Process at http://ec.europa.eu/education/policies/educ/bologna/bologna_en.html.
[23] See ERA Green Paper, http//:ec.europa.eu/research/era/pdf/era-greenpaper_en.pdf, and "Euroscience on ERA", http://cordis.europa.eu/search/index.cfm?fuseaction=news.simpledocument&N_RCN=28441
[24] See European Strategy Forum on Research Infrastructures homepage at http://cordis.europa.eu/esfri/
[25] Katy Borner discusses this problem, emphasising the need for computational structures that can "support the 'global brain' that is emerging on the planet". Katy Borner, Associate Professor of Information Science, Indiana University, http://ella.slis.indiana.edu/~katy/index.html

[26] Jose Manuel Durao Barroso, President of the European Commission, "ICT industry has a major role to play in the European economy of the 21st Century", CeBIT Trade Fair, Hannover, 3 March, 2008.
[27] See EC Enterprise and Industry, "e-Skills for the 21st Century: fostering competitiveness, growth and jobs" at http://ec.europa.eu/enterprise/ict/policy/ict-skills.htm
[28] See EC COM (2007) 496 final, e-Skills for the 21st Century: fostering competitiveness, growth and jobs.
[29] ACM Curricula Recommendations at http://www.acm.org/education/curricula.html
[30] See Jeannette M. Wing, "Computational Thinking" in Communications of the ACM, Viewpoint, March 2006/Vol. 49, No. 3, pp. 33-35.
[31] See INSPIRE (Infrastructure for Spatial Information in Europe) homepage at http://www.ec-gis.org/inspire/
[32] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html
[33] 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.
[34] See SURA Grid Technology Cookbook at http://www.sura.org/cookbook/gtcb/
[35] 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
[36] 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.
[37] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html
[38] 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/
[39] 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.
[40] 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
[41] 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
[42] 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

[43] See Foundation for Information Policy Research, Text of Directive 2001/29/EC at http://www.fipr.org/copyright/eucd.html
[44] 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
[45] ICEAGE: www.iceage-eu.org/library and EGEE: http://egee.lib.ed.ac.uk
[46] See OGF wiki, "Towards Professional Grid Certification" draft document at http://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/TowardsProfessionalGridCertification
[47] 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
[48] 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
[47] Please see the EC European Qualifications Framework page at http://ec.europa.eu/education/policies/educ/eqf/index_en.html.
[49] See EGI website for current information on NGI development in each Member State:
[50] See EGI Knowledge Base Main Page at http://knowledge.eu-egi.eu/index.php/Main_Page
[51] See ESFRI Roadmap at http://cordis.europa.eu/esfri/roadmap.htm
[52]See ICEAGE Forum Agenda Geneva, Third Forum Meeting, EGEE Conference at http://www.iceage-eu.org/events/forumMeetingGeneva.html
[53] 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/
[54] See ICEAGE Forum Agenda Washington, Second Forum Meeting Notes, 14 September 2006, at http://www.iceage-eu.org/events/forumMeetingWashington.html
[55] Future promotion of grid education may need to take into account the development of the European Institute of Technology (EIT), which would aim to foster integration of the knowledge triangle across the EU. The EIT could play a part in standardisation (accreditation and policy for embedding of e-Infrastructure in educational establishments across Member States, for instance). For more information please see EC COM (2006) 0604 – C6-0355/2006/0197 (COD), http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A6-2007-0293&language=EN&mode=XML, and "Parliament Backs European Institute of Technology", Euractiv, 26 September, 2007, http://www.euractiv.com/en/science/parliament-backs-european-institute-of-technology/article-167110
[56] See "Towards Professional Grid Certification", draft doc 14419, ET-CG GridForge, OGF at https://forge.gridforum.org/sf/go/doc14419?nav=1
[57] 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