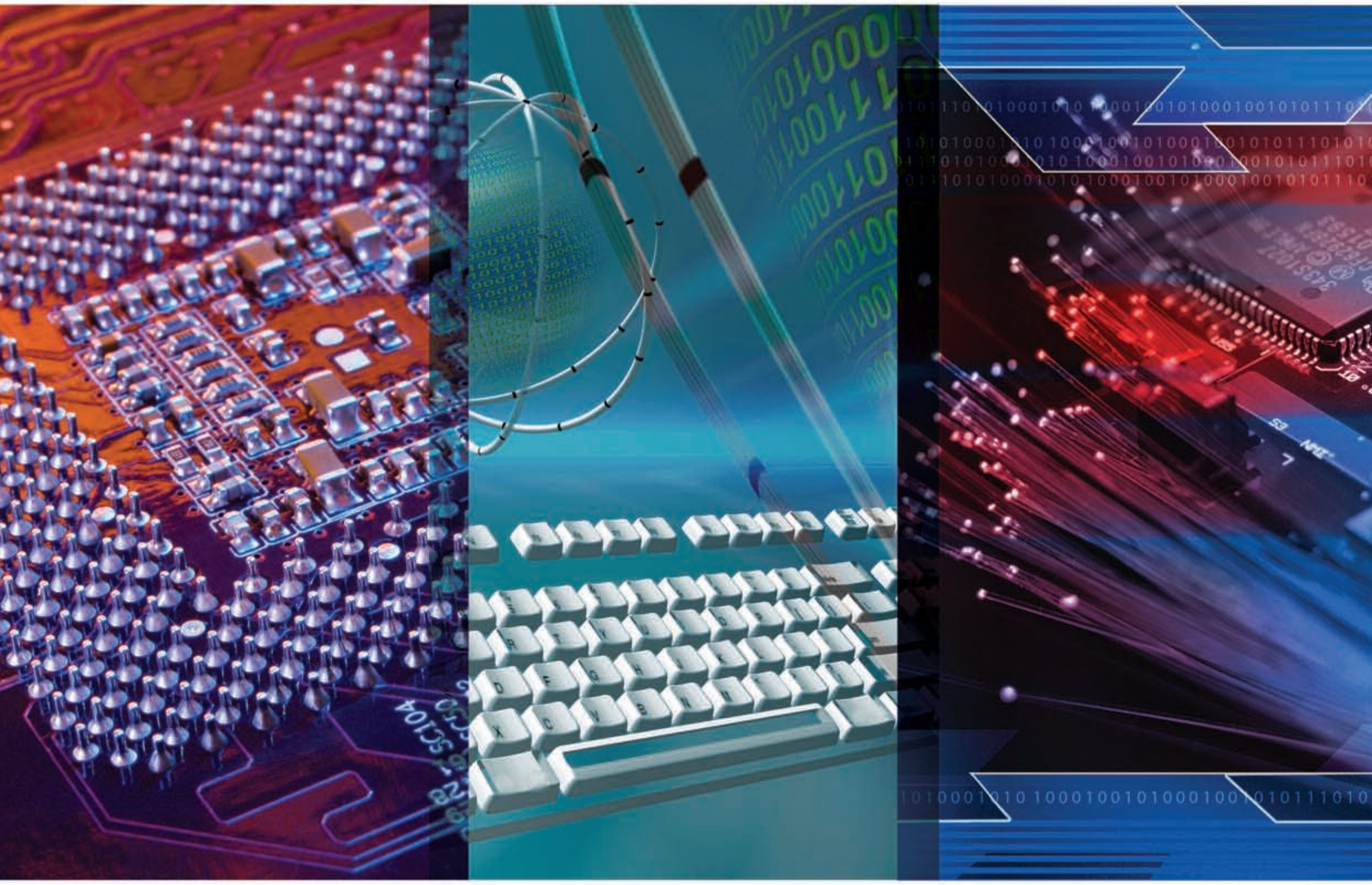
Title
RCUK Review of e-Science 2009: Building a UK foundation for the transformative enhancement of research innovation

Authors
Atkins, D. E.
Borgman, Christine L.
Bindhoff, N.
et al.

Publication Date
2010-05-21

RCUK Review of e-Science 2009

BUILDING A UK FOUNDATION FOR THE TRANSFORMATIVE ENHANCEMENT OF RESEARCH AND INNOVATION



RESEARCH COUNCILS UK

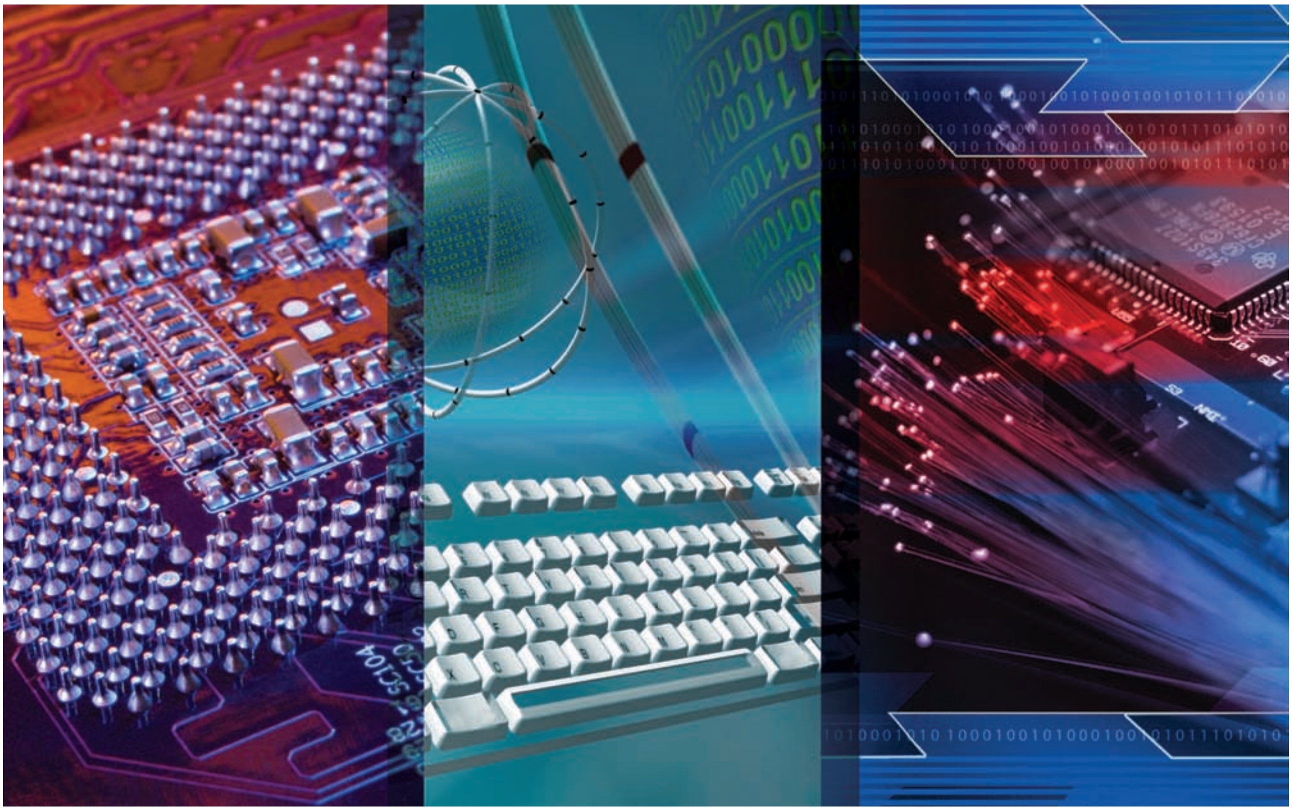
THE ROYAL SOCIETY

Report of the International Panel for the 2009 Review of the UK Research Councils e-Science Programme

International Panel for the 2009 RCUK Review of the e-Science Programme

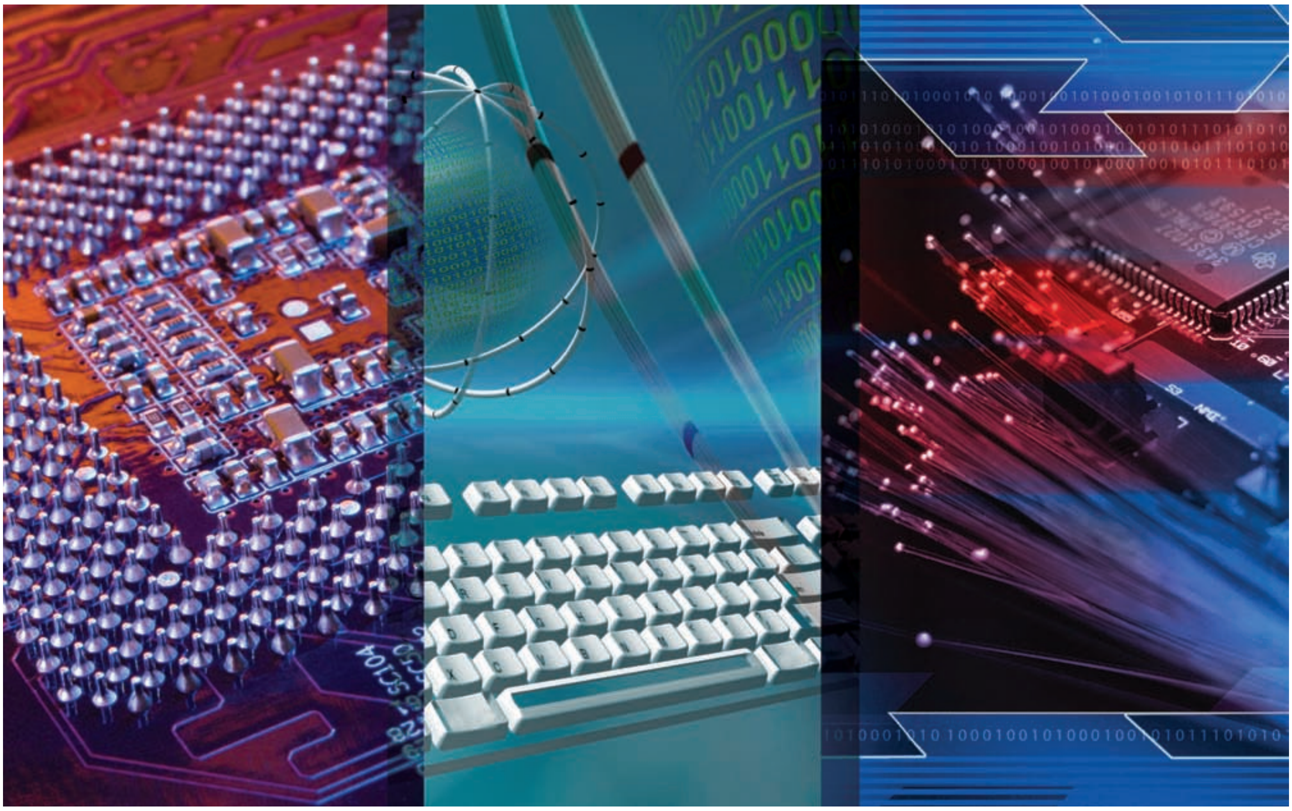
Anders Ynnerman (Linköping University, Sweden) Paul Tackley (ETH Zürich, Switzerland) Albert Heck (Utrecht University, Netherlands) Dieter Heermann (University of Heidelberg, Germany) Ian Foster (ANL and University of Chicago, USA) Mark Ellisman (University of California, San Diego, USA) Wolfgang von Rüden (CERN, Switzerland) Christine Borgman (University of California, Los Angeles, USA) Daniel Atkins (University of Michigan, USA) Alexander Szalay (John Hopkins University, USA) Julia Lane (US National Science Foundation) Nathan Bindoff (University of Tasmania, Australia) Stuart Feldman (Google, USA) Han Wensink (ARGOSS, The Netherlands) Jayanth Paraki (Omega Associates, India) Luciano Milanesi (National Research Council, Italy)

Table of Contents



Annexes

Executive Summary



About this Report

This is the report of an international Panel of 16 experts convened by an RCUK Steering Committee on behalf of all UK Research Councils to review the RCUK e-Science Programme. The Review focussed on e-Science as a whole, across all Research Councils' remits. It focussed on the added value from the Programme as a whole rather than separately reviewing each individual element. The Steering Committee in concert with the Chair also (1) agreed on a set of questions to guide the assessment process termed the "e-Science Evidence Framework"; (2) selected the 63 projects from 24 institutions to be visited by the Panel; (3) selected background data to be provided to the Panel; (4) created the agenda for the review week; (5) received a preliminary briefing from the Panel at the end of the review week; and (6) orchestrated the disposition and use of this final report.

Background: e-Science and the Core Programme

In 2001, the UK Research Councils launched one of the earliest multidisciplinary "e-Science" initiatives, recognising that the push of technology – the enormous and growing capacity of computing, storage, communication and software systems – offered the opportunity not only to automate science but also to apply new methods that could revolutionise how science was performed. e-Science was defined as "research done through distributed global collaborations enabled by the Internet, using very large data collections, terascale computing resources and high performance visualisation." The Programme was ring-fence funded, allocated between an "e-Science Core" – a matrix structure spanning all Councils – and the disciplinary Councils. All ring-fence funding was intended to be used in a coordinated, leveraged way to enhance both the infrastructure for e-Science (e-infrastructure) as well as the meaningful application of e-Science methods to high quality research. Although some support for core activities was continued after 2006 the dedicated Core Programme was phased out after 2006, and the Programme, although it is intended to still have Council-wide scope, does not enjoy ring-fence funding. Over the past eight years, the funding trajectory and the management structure for the Programme have changed along with national research priorities, funding organisations and assessment structure. These changes could be used in concert to further strengthen an ongoing strategic e-Science Programme.

Current Status of the Programme and its Impacts

The inaugural goals of the e-Science Core Programme (eSCP) were to assist the development of essential, well engineered, generic Grid middleware usable to both e-Scientists and industry; provide necessary infrastructure support for UK e-Science projects; collaborate with the international e-Science and Grid communities; and develop a framework in collaboration with scientists, computer scientists and industry to promote the emergence of robust, industrial-strength Grid middleware. The complementary goals of the individual Research Councils' e-Science investments were to support and leverage the Core Programme investments in service of their disciplinary and interdisciplinary mission of research excellence, impact, public engagement, training and provision of world-class research facilities.

The most obvious and expected impacts of the Programme are in academia. Perhaps the largest and most important academic impact is in the interdisciplinary efforts that have gone into most of the projects and that have fostered new social academic networks and spawned new collaborations. The Programme was a forerunner to many other similar programmes worldwide, put the UK at the forefront of the development of e-Science, and contributed to the acceptance of e-Science as a valid research paradigm.

e-Science serves as a platform for research enhancement, including systems (shared software with interfaces and a variety of extensions features), organisational structures (formal and informal groups that provide very important e-Science services), human capital (build-up of knowledge and experience that makes it much easier to adapt and adopt technologies), and data and information resources (systems that support the increasingly data-intensive nature of research):

- The Programme has been successful at creating shared software tools, ontologies and standards work. It is important to continue to sustain long-term centres for such developments.

- Hardware investments have likewise grown at regional and university facilities, connected by grids, but these are no longer impressive national resources and should be re-examined for alternative, cost-effective ways of providing superior storage and computational power.

- Investments in human capital have provided people with training and experience in cross-disciplinary and computational thinking, but the current risk is that, without future Programme funding, these young people will face uncertain careers and may be lured to industry and academia outside the UK. Even so, some of the community's organisations continue to provide education, training and annual meetings.

- In the area of data and information stewardship, the Programme has supported relevant research and tool building to provide access to information but needs more extensive planning for large-scale support of storage or federation of data resources as data grows exponentially.

- Some heavily used services created and supported by the Programme include the National Grid Service, Access Grid and the Digital Curation Centre; however, there is no plan for long-term basic support outside of R&D project budgets, so we advise revisiting the decisions on which services need to be provided specially or can be procured commercially, and then making a strategic plan in collaboration with the research community, service industry and government.

- Important e-Science research facilities that have extra-ordinary levels of utilisation by their respective communities and high status within their communities need to be provided with continuous, sustained funding, but as they do not themselves produce traditional research outcomes, they are not appropriately funded through normal research project competitive processes.

The UK's investments have clearly had a positive effect, in particular through direct links with industry, in three important areas – life sciences and medicine, materials, and energy and sustainability – and important secondary economic impacts on health and social wellbeing.

Future Considerations

Many technologies funded by the Programme are already part of the research modus operandi for many scholars and they have increased productivity and capability enormously. They promise even greater value as they spread throughout the community – but only if there is a credible plan for long-term support so that researchers can count on their availability. The greatest risk of not "crossing the chasm" from early adopters to the general population of researchers is that these technologies will not realise their early promise or pay back the investments. It is essential to make the necessary organisational and financial investments to support this transition. Some areas requiring sustained resources include superior networking; distributed computing; Grid architecture; data curation, sharing and management; storage management, semantic technology, process and workflow management, and Web technology.

Our forecast for the next five years is movement from research-level developments to sustained infrastructure. This includes maintaining critical-size centres that are already established, dramatically expanding community involvement, providing mid-term career paths for current personnel, building stronger bridges to the high performance computing (HPC) community, and developing shared infrastructure that is reliable, mature and sustainable. It is also important to sustain and grow the emerging community through strong national leadership, stable cross-council coordination, community building and training, systematic dissemination of best practices and service-based e-Science.

Our vision for the next ten years requires addressing the challenges of transformations in technologies, in disciplines and in the use of data. We expect that different scientific disciplines will continue to adopt e-Science technologies at a different rate. The current hierarchy of centres and resources allows a more gradual evolution of the system and support infrastructure and should be maintained. Data-intensive computational technologies are evolving extremely rapidly, and establishing a shared, national support centre to take an exploratory role would be a very cost-efficient way to approach this inevitable shift.

The Panel suggests that the established platform of e-Science not only could scale to broader use, but also could move to more transformative impact. The UK's pioneering leadership in e-Science continues to require nurturing and special handling, for it has barely begun to achieve its full potential. We would even suggest that the UK has a responsibility to the global science community to continue its leadership in e-Science to produce vision, models, best practices, and competition to inform and motivate others, and we propose a departmental structure that could launch and support an even bolder e-Science, e-research, e-learning initiative in ways that would contribute to building the nation's future economic strength.

Responses to Evidence Framework Questions

The following are very brief responses to the questions in the Evidence Framework.

1. Did the UK e-Science Programme build a Platform which enables e-Science tools, infrastructure and practises to become incorporated into mainstream research in the UK?

There is evidence that significant capability has been created, both in tools and practice. The adoption varies widely across disciplines and projects, and sustainability is a concern. Even so, we have seen numerous examples of e-Science facilitating flows of ideas across disciplinary fields, supporting new ways to extract knowledge from large data collections, augmenting discovery with computational methods, and grounding and informing advances in computer science and engineering. Also, there are still not enough regular course offerings, high-quality course materials and education initiatives at the undergraduate and post-graduate level.

2. How does UK e-Science activity compare globally?

The Programme is in a global leadership position in scientific data management, workflow environments and Grid architecture deployment. Their international engagements around e-Science have been in general "best with best," leading in a broad range of areas, but the Programme as a whole has yet to create the right environment for sustaining important research facilities. The Programme attracted new scientists looking to combine e-Science with their application domain, but the capacity of the Programme to nurture and support e-Science researchers at every stage of their career has been much more limited.

3. What has been the impact (accomplished and potential) of the UK e-Science Programme?

A surprisingly large fraction of the presented projects had in one way or another generated recognisable commercial impact. Also, the UK has built a significant and diverse portfolio of important projects addressing major technical and societal challenges in the physical, biological, social and information sciences, with the potential to spur practical improvements in health care and the environment.

4. What are the future opportunities for UK e-Science?

Moving forward requires coordination, clever design, effective leadership, and long-term commitment to a system of linked and balanced interaction between the various communities and sponsors. Processes need to be established to guide and sustain the creation of a shared e-infrastructure on which project-specific software can be easily tailored. We recognise that funding for research is today a huge challenge, but perhaps the UK can take the opportunity for funding reallocation into an even bolder strategic e-Science activity with potentially greater societal benefit across all science fields and education at large.

5. How did the Programme Strategy (having a Core and individual Research Council Programmes, developing tools and applications in parallel) affect the outputs from UK e-Science?

The organisational and management strategy has worked well; however, the Programme should not be viewed as a one-shot, five-year project, but rather as a long-term strategy that needs to be continuously refined and carried forward indefinitely as a more permanent crosscutting Programme with real authority and resources. We suggest a hierarchy of small but energetic, value-adding coordinating and leveraging organisations for e-Science.

Major Conclusions and Recommendations

The Panel has concluded that the UK e-Science Programme is in a world-leading position along the path of Building a UK Foundation for the Transformative Enhancement of Research and Innovation. The UK has created a "jewel" – a pioneering, vital activity of enormous strategic importance to the pursuit of scientific knowledge and the support of allied learning. Investments are already empowering significant contributions to well-being in the UK and beyond. The UK must decide whether to create the necessary combination of financial, organisational and policy commitments to capitalise on their prior investments, and to move to the next phase of building capability, growing adoption and achieving competitive advantage.

The successful creation and adoption of e-Science is an organic, emergent process requiring ongoing, coordinated investment from multiple funders together with coordinated action from multiple research and infrastructure communities. It requires nurturing robust infrastructure in a cycle that couples research, application development and training processes. The balance between these processes drives success. None of this is easy, but the rewards for success are enormous.

We conclude this report with a list of a dozen major recommendations for action at a more general level than in earlier sections. The emphasis here is on what to do, rather than being prescriptive about how to do it. These are not in a priority order, but most are quite interdependent.

1. Establish organisation and management structures that continue to treat e-Science as a designated strategic initiative spanning all Research Councils and having ongoing designated funding.

2. Establish more systematic and better supported mechanisms, including targeted funding, to nurture collaboration and bi-directional knowledge transfer between academia and industry in the creation, provisioning and application of e-Science.

3. Sustain and strengthen the RCUK network of e-Science Centres.

4. Sustain the operational e-infrastructure for e-Science created to the present.

5. Recognise in Programme calls and funding policies that there are people in several complementary roles that need to be funded in a balanced way.

6. Continue funding policies that strongly encourage or require the creation and adoption of shared e-infrastructure.

7. Encourage and support even more participation of the arts and humanities research communities in the e-Science Programme.

8. Encourage and support even greater leadership by the social science research community in the adoption of e-Science methods.

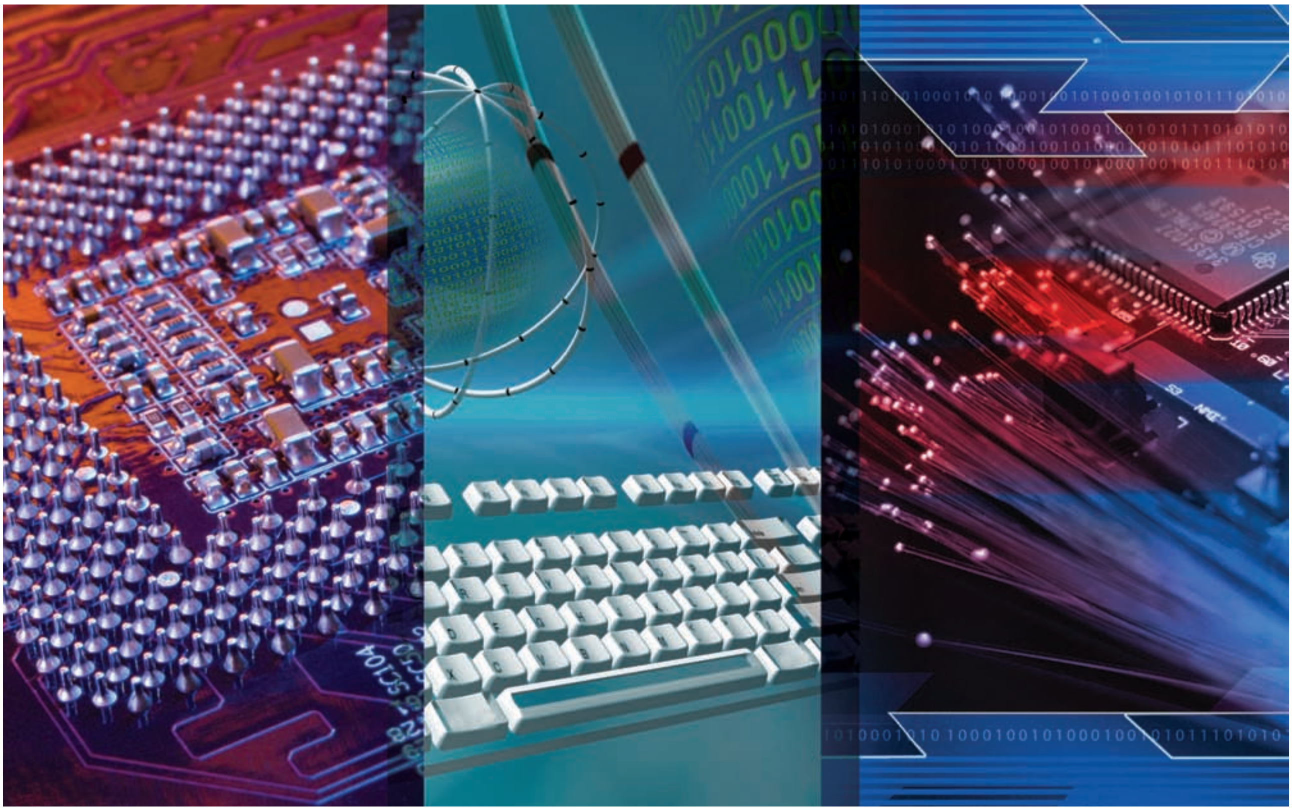
9. Develop a dual strategy that both (1) accelerates the adoption of e-Science methods in the "mainstream market" of researchers; and (2) refreshes the investments in the "early market" to produce the next wave of innovation in e-Science services and application.

10. Continue the strong focus on creating practices and services for appraisal, curation, federation and long-term access to scientific data.

11. At every opportunity establish and support policies for openness: open-source code, open data and open courseware.

12. Place greater emphasis on the overarching goal of establishing capacity for collaborative, international, interdisciplinary team science to occur routinely on IT-enabled platforms provide all the services required by the research teams, and enable them to work together well in all four variations of same and different, time and space.

Table of Acronyms
Acknowledgements



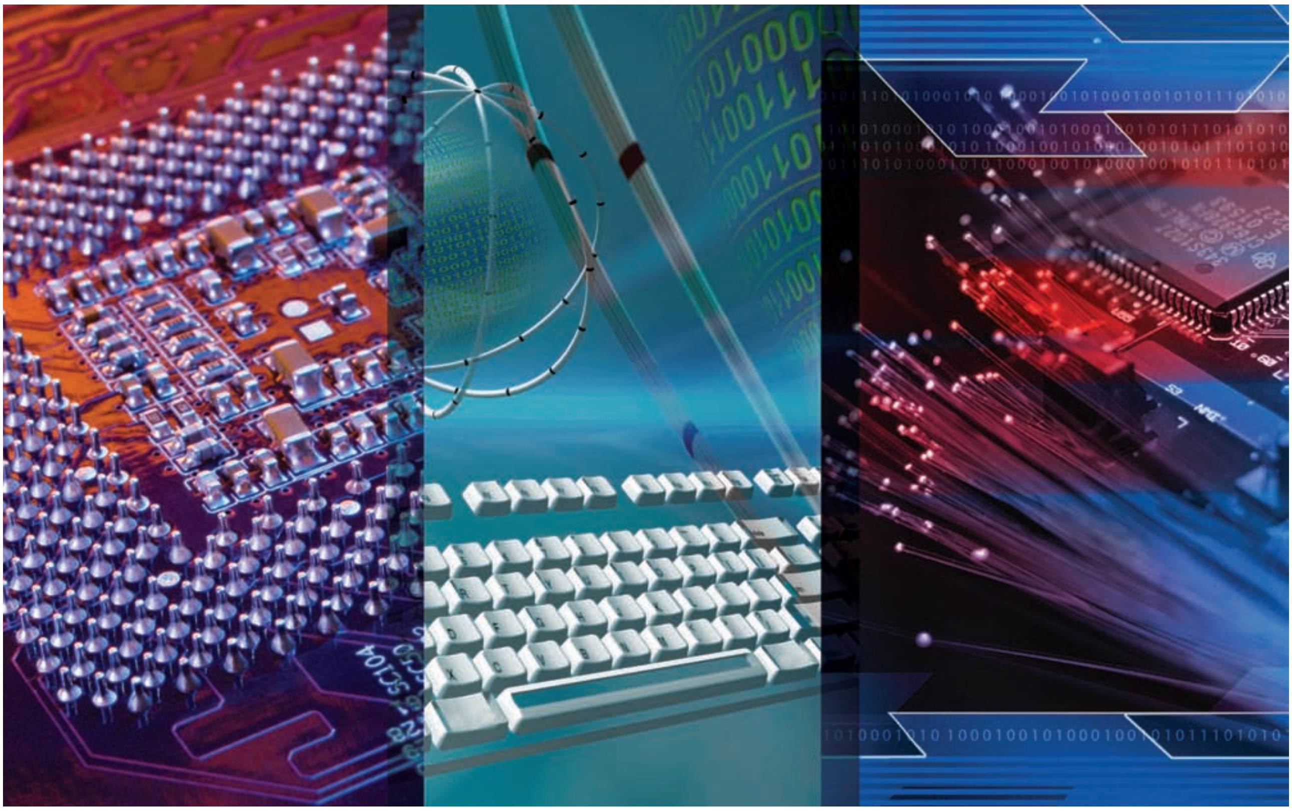
AHM	All Hands Meeting	
AHRC	Arts and Humanities Research Council	
BBSRC	Biotechnology and Biological Sciences Research Council	
BEP	Bioinformatics E-Science Programme	Referred to as BEP1 and BEP2
BERR	Department for Business, Enterprise and Regulatory Reform	In 2009, BERR merged with another Government department (DIUS) to form BIS (Department for Business, Innovation and Skills)
BIS	Department for Business, Innovation and Skills (established 2009)	Formerly DIUS and BERR
CCLRC	Council for the Central Laboratory of the Research Councils	Merged with PPARC to become STFC (Science and Technology Facilities Council)
CERN	European Organisation for Nuclear Research	
Co-I	Co-investigator	
CSA	Chief Scientific Advisor	
CSR	Comprehensive Spending Review	
DELNI	Department of Education and Learning in Northern Ireland	
DGSI	Director General of Science and Innovation	
DIUS	Department for Innovation, Universities and Skills	Formerly OSI. In 2009, DIUS merged with another Government department (BERR) to form BIS (Department for Business, Innovation and Skills)
DTA	Doctoral Training Account	
DTC	Doctoral Training Centre	
DTI	Department for Trade and Industry (abolished 2007)	When the DTI closed many of its functions were transferred to BERR
EPSRC	Engineering and Physical Sciences Research Council	
ESRC	Economic and Social Research Council	
EU	European Union	
fEC	Full Economic Costing	
FTE	Full-Time Equivalent	
HE/HEI	Higher Education/Higher Education Institution	
HECToR	High End Computing Terascale Resources	
HEFCE	Higher Education Funding Council for England	
HEFCW	Higher Education Funding Council for Wales	
HESA	Higher Education Statistics Agency	

HPC	High Performance Computing	
ICT	Information & Communications Technology	
ILL	Institut Laue-Langevin (France)	
IP/IPR	Intellectual Property/Intellectual Property Rights	
JISC	Joint Information Systems Committee	
KT/KTN	Knowledge Transfer/Knowledge Transfer Network	
KTP	Knowledge Transfer Partnership	
LHC	Large Hadron Collider	
MOD	Ministry of Defence	
MRC	Medical Research Council	
NCeSS	National Centre for e-Social Science	
NERC	Natural Environment Research Council	
NSF	National Science Foundation (US counterpart to UK Research Councils)	
OSI	Office of Science and Innovation	OSI was a Government department which was formerly known as OST (Office of Science and Technology) and became DIUS (Department for Innovation, Universities and Skills)
OST	Office of Science and Technology	OST was established in 2007 and then became the OSI
PI	Principal Investigator	
PPARC	Particle Physics and Astronomy Research Council	Merged with CCLRC to form STFC (Science and Technology Facilities Council)
R&D	Research and Development	
RA	Research Assistant	
RAE	Research Assessment Exercise	
RAL	Rutherford Appleton Laboratory	
RCUK	Research Councils United Kingdom	
RDA	Regional Development Agency	
SFC	Scottish Funding Council	
SME	Small/Medium sized Enterprise	
SR	Spending Review	
SRIF	Science Research Investment Fund	
STFC	Science and Technology Facilities Council	Formerly PPARC and CCLRC

Panel members are grateful to the Steering Committee for their excellent preparatory work that led to a well-structured, albeit busy schedule. The Panel also offers an enormous vote of thanks to the EPSRC staff involved in managing the review, and to the other Research Council staff who assisted and provided information throughout the review process. Specifically, their tireless efforts in ensuring members of the Panel were in the right place at the right time, their overall project management and preparation, including distribution of documentation for the Review, and first-rate travel arrangements contributed enormously to a memorable experience and the success of the Review.

The Panel is deeply appreciative of the time and effort given to the review by the institutions and their staff. Each day, the Panel was exposed to presentations of high quality. The Panel appreciates the enormous time and effort devoted to the RCUK Review by the e-Science community including their rapid response to provide information we sought during the review week. Talks and the interactions with the Panel formed the basis of lively discussions and an effective communication vehicle. Many academic researchers and industrialists travelled significant distances to participate in the presentations and open dialogue with Panel members. Here again, the Panel is grateful for the time and effort generously given over to the Review by all those involved. The many frank discussions with academics and industrialists were invaluable and form the basis of the Panel's findings and recommendations. The Panel is especially appreciative of the early career researchers that openly shared their diverse experiences.

1. Introduction and Background



Purpose of the Review

This is the report of an international Panel of 16 experts convened by an RCUK Steering Committee on behalf of all UK Research Councils to review the RCUK e-Science Programme begun in 2001 and operating with ring-fenced funding and a core programme structure through 2006. From 2006 to date, e-Science funding has operated without an explicit core and all Research Council funding through individual Councils. The Steering Committee and review process were supported and staffed by the Engineering and Physical Sciences Research Council (EPSRC). The RCUK Review of e-Science 2009 focussed on e-Science as a whole, across all Research Councils' remits. It focussed on the added value from the Programme as a whole rather than separately reviewing each individual element. The purpose of the Review - the Terms of Reference - is to:

- Assess the impact of the Programme on research areas nationally and internationally; on broader wealth creation and quality of life; and on e-Science itself;
- Assess and compare the quality of the UK research base in e-Science with the rest of the world via triangulation of data, Panel and community perception;
- Comment on the added value of this Programme;
- Present findings and recommendations about the strength, weakness and opportunities for the future to the research community and Councils.

Structure of the Review

The review process oversight was the responsibility of a Steering Committee chaired by Dr David Snelling of Fujitsu Laboratories of Europe Ltd. The full membership is given in Figure 1.

This Steering Committee, after receiving nominations from the community, selected the Panel Chair, Professor Daniel E. Atkins, Professor of Information, Computer Science, and Electrical Engineering at the University of Michigan, U.S.A. and the inaugural Director of the Office of Cyberinfrastructure at the U.S. National Science Foundation. In concert with the Chair, the Steering Committee selected the other 15 members - all based outside the UK. The Panel members represented a balanced mix of international researchers spanning the RCUK disciplines and with extensive experience in e-Science – the enhancement of scientific research based upon platforms of advanced digital computing and communication technology aka, e-infrastructure. Brief academic biographies of the Panel members are included in Annex A.

The Steering Committee in concert with the Chair also (1) agreed on a set of questions to guide the assessment process termed the "e-Science Evidence Framework"; (2) selected the 63 projects from 24 institutions to be visited by the Panel; (3) selected background data to be provided to the Panel; (4) created the agenda for the review week; (5) received a preliminary briefing from the Panel at the end of the review week; and (6) orchestrated the disposition and use of this final report. EPSRC provided Review Secretariat support for the organisation, planning and logistics.

RCUK Review of e-Science Steering Committee

- Dr David Snelling (Chair)
  Fujitsu Laboratories of Europe Ltd
- Professor Geoff Boxshall
  The Royal Society
- Dr Lesley Thompson
  Director Research Base, Engineering and Physical Sciences Research Council (EPSRC)
- Ms Suzanne Mills
  Senior Policy Manager, Economic and Social Research Council (ESRC)
- Dr Janet Seed
  Associate Director, Science Programmes, Science and Technology Facilities Council (STFC)
- Dr Amanda Collis
  Head Engineering & Biological Systems, Biotechnology and Biological Sciences Research Council (BBSRC)

Figure 1 - Steering Committee membership

The five top-level questions in the Evidence Framework are as follows:

1. Did the UK e-Science Programme build a Platform which enables e-Science tools, infrastructure and practises to become incorporated into mainstream research in the UK?

2. How does UK e-Science activity compare globally?

3. What has been the impact (accomplished and potential) of the UK e-Science Programme?

4. What are the future opportunities for UK e-Science?

5. How did the Programme Strategy (having a Core and individual Research Council Programmes, developing tools and applications in parallel) affect the outputs from UK e-Science?

The framework was used extensively by the Panel in the review process described below. These questions are not a template for the organisation of this report, but responses to all questions are embedded in the report and briefly summarised in Section 4 within the structure of the Evidence Framework.

Review Panel Activities

The full Panel convened in and near Oxford on December 6th through to the 11th for an intense week of review. The Review Secretariat, under the general supervision of the Steering Committee, provided a 200 page Evidence Document describing the UK e-Science Programme in terms of its people, funding, organisation and policy. This overview document was compiled from the Research Councils' management records and the UK e-Science research community. The documentation included (1) descriptions on the structure, history and spending trajectories; (2) summaries of funded projects; and (3) summaries of surveys sent to grantees. The data provided for use by the Review Panel will be published on the RCUK web site on completion of the Review (documentation provided in confidence will not be published). The Panel requested some additional information from the Secretariat as well as from grantees during the review week.

Especially noteworthy in the evidence documents, besides the compilation of relevant data, is the extensive narrative provided by the e-Science research community itself on the questions in the e-Science Evidence Framework prepared by the Steering Committee. This was essentially a self-assessment but, as should be clear in this report, it is to a very large extent confirmed by the external Panel. The Panel appreciates the enormous amount of effort on the part of the community and the Secretariat that went into gathering and compiling this document.

In a mixture of full and sub-panel formats, the Panel interacted with researchers from about 60 funded e-Science projects either at the All Hands e-Science Meeting, the IEEE e-Science Meeting, or on visits to universities within a 2-hour drive from Oxford. The majority of the interactions occurred in the conference facilities of the Oxford Kassam Football Stadium. As the Panel gathered input, the Panel's sub-groups corresponding to the primary questions of the Evidence Framework, processed what we heard, read and observed against those five areas and also add additional topics as appropriate.

All members of the Panel were given the extensive Evidence Document shortly before the meeting in Oxford. The Panel members skimmed the material but conducted the review week largely independent of deep knowledge of the material in the document. This provided an additional, more independent assessment during the week, and was in part a practical necessity given the intensity of the Panel's schedule. Subsequent to the review week, the Panel, and particularly the Chair, have studied the Evidence Document closely to enable drawing upon it in this report, and to comment on the extent of our concurrence with the community self-assessment presented in Section 4 of the Evidence Document. We assume that interested readers of this report will have access to the RCUK-produced Evidence Document and we will not attempt to reproduce the material in it in this report. We will, however, refer to evidence in the report as part of our findings and recommendations.

As the Panel read in the Evidence Document and heard more about in the review meetings, the Department for Trade and Industry (DTI) (subsequently replaced in the BIS formation) played a significant role in co-funding the early stages of the e-Science Programme including funding of projects not funded by the Research Councils. Although we were not asked to review these projects, our general sense is that the inclusion of industrial perspectives and economic innovation goals in the overall e-science Programme has been an important component of its success. As we mention later, this success could be further amplified within the context of the current BIS structure.

In closed meetings the Panel achieved collective understanding about the context, vision and opportunities afforded by the e-Science Programme; made judgements on strengths and weaknesses of the

Programme to date; and developed recommendations about the opportunities for the future. After the meeting in Oxford, the five subgroups prepared written responses to the questions and any additional material they deemed important and sent it to the Panel chair for integration. The Chair prepared the draft report which was refined with further input from the Panel to produce this report sent to the RCUK. The draft report was circulated for comment by the UK e-Science community prior to a Town Meeting in London on 9th February 2010. At this meeting the Chair made an oral presentation about the Panel's report and invited written comment from the community. All input was processed by the Chair to produce this final report.

Attribution of Credit

In this report, the Panel has chosen to mention by name the role of some individuals based on our readings, presentations and conversations. We readily acknowledge that this is a risky undertaking in that, without doubt, we have failed to mention others to whom the Panel were not exposed, but who played important roles in the e-Science Programme. We apologise to any who feel slighted. We realise that this was a team effort of extraordinary proportions.

Background on the UK e-Science Programme

Information and computing technology (ICT) has, since the inception of the modern digital computing age, been both a product of and a major tool for scientific and engineering research. The application of ICT to research has pioneered ICT development and adoption in many other fields: commerce, defence, entertainment and general social well-being. This opportunity continues in the current era of e-Science. ICT has evolved in raw power and reach at exponential rates for about 60 years and is now the basis for ubiquitous environments to support humans in a huge range of discovery, learning, entertainment and social engagement activities.

The use of computers for modelling, simulation and prediction is well established under the phrase *computational science* and is now widely regarded as a third leg of scientific method together with theory and traditional experimentation. The increasing power of computers (tera- to peta- to exa- scale) is offering revolutionary opportunities to tackle understanding of extremely large and complex systems involving multiple scales of time and distance (multi-scale) and models and knowledge from many fields (multi-science). Significant effort is still required to seize this opportunity.

In the 1990s a growing number of research communities and projects around the world realised that the push of technology – the enormous and growing capacity of computing, storage, communication and software systems – offered the opportunity not only to automate science, but to apply new methods that had the potential to revolutionise how science was performed[1]. This could enable the tackling of grand challenge problems not possible in any other way. The opportunity is not limited to computation. It includes requirements for stewardship, curation and mining of enormous collections of heterogeneous data; online observatories and research instrumentation; and new types of virtual research environments. Virtual research environments (VRE) or "collaboratories" enable broadened participation in science and can dramatically reduce barriers of time and distance (distance in the geographic, disciplinary and organisational sense). Creating a successful VRE requires a sophisticated socio-technical approach to the design and adoption of the system.

One such community was the global particle physics community mobilised around the LHC at CERN. The LHC community faced enormous challenges due to the scale of data produced (petabytes), the highly distributed nature of their community and the enormous complexity of the central instruments they were building. The solution was a global grid of computing and data storage resources and they sought support from funding agencies all over the world.

The UK Research Councils under the leadership of Director General Sir John Taylor responded to the UK LHC community, but fortunately not in a one-off way. Realising the emerging inflection point in the broad use of ICT for research, he rather used it as an occasion to launch in 2001 one of the earliest multidisciplinary "e-Science" initiatives. e-Science was defined as "research done through distributed global

[1] The antecedent of the "e" in e-Science is rarely made explicit. It no doubt began as "electronic" but now better stands for "enhanced" or "enabled."

collaborations enabled by the Internet, using very large data collections, terascale computing resources and high performance visualisation." This definition was often paired with a definition of an infrastructure enabler for e-Science, the *Grid*: "a new generation information utility requiring middleware, software and hardware to access, process, communicate and store huge amounts of data." This definition includes mention of high-performance computing (terascale at the time), but focuses on collaboration, data and a Grid architecture platform.

More generally, the initiation of e-Science programmes in the UK and elsewhere, is driven by the growing conviction that e-Science is now essential for meeting 21st century challenges in scientific discovery and learning due to at least the following trends in scientific research:

- the multi-scale and multi-science nature of today's frontier science challenges;
- the requirement for a multidisciplinary, multi-investigator, multi-institutional approach, often international in scope;
- the large and growing data intensity and heterogeneity from simulations, digital instruments, sensor nets and observatories;
- the increased scale and value of data and demand for semantic federation, active curation and long-term preservation of access; and
- the need to engage more students in high-quality, authentic, passion-building science and engineering education.

An e-Science Programme itself is a complex emergent system with multiple coupled objectives, many stakeholders and actors, and the need for skilful balancing acts. It is both a platform for research as well as an object of research. To give the new Programme a fighting chance, the RCUK established a new Programme with a then new type of alignment between mission, authority and resources. It appointed Professor Tony Hey as the Director and established an effective steering and management structure including representatives from the Research Councils, industry, international communities and e-infrastructure operational groups. The Programme was ring-fenced funded (£213M over 5 years from RCUK), allocated between an "e-Science core" - a matrix structure spanning all Councils - and the various disciplinary Councils. All ring-fence funding was intended to be used in a coordinated, leveraged way to enhance both the infrastructure for e-Science (e-infrastructure) as well as the meaningful application of e-Science methods to high quality research.

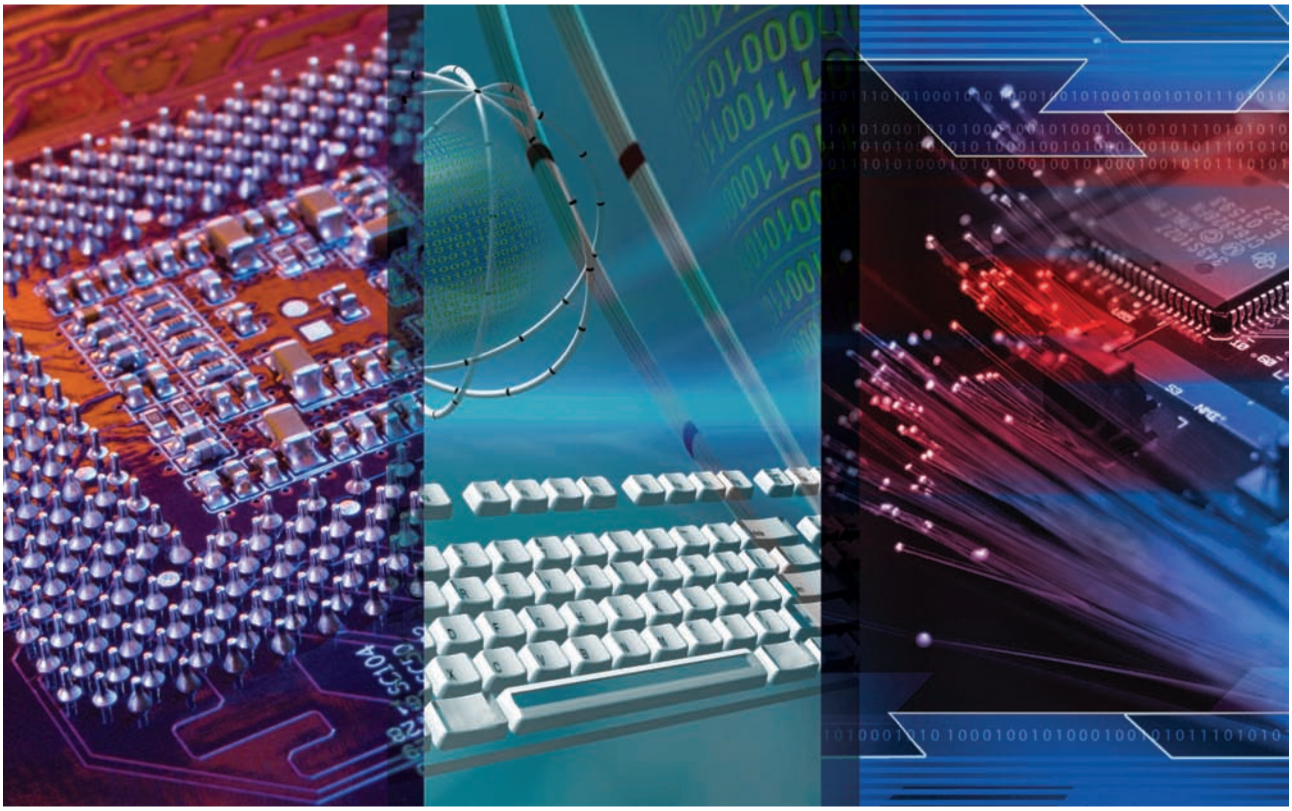
The funding in the Core was intended to leverage and steer the funding in the Councils towards shared e-infrastructure, interoperability, transfer of best practices, support of interdisciplinary, knowledge transfer to industry, etc. The Department for Trade and Industry (abolished through restructuring in 2007) also contributed funding to help create incentives for industrial participation and provide additional funding flexibility. For the five years of the initial Core Programme, including both ring-fenced Council and DTI funding, about 21% of the funding was distributed through the Core and the remainder through the Councils. A similar balance (about 25% core) exists in the US National Science Foundation Office of Cyberinfrastructure (OCI), analogous to the Core, and the various research Directorates, analogous to the Councils, although the remits are not exactly equivalent.

The Core Programme as originally conceived and directed by Professor Hey was phased out after 2006. The e-Science Programme, although it is intended to still have Council-wide scope, does not enjoy fenced funding. It is now managed by the EPSRC with a similar set of advisory groups. An e-Science Envoy, Professor Malcolm Atkinson, has been appointed to advise and support the EPSRC in the nurturing of the e-Science Programme and to serve as an ambassador for e-Science nationally and internationally. The Envoy does not have the mission, authority, nor the resources previously afforded Professor Hey. We understand that there is still funding going into e-Science but not necessarily in ways that it is tagged as e-Science and so we could not get a clear picture of how much funding is currently allocated to e-science. Both JISC and the Science and Technology Facilities Council are also investing in e-infrastructure to underpin e-Science.

Over the past eight years, both the funding trajectory and the management structure for the e-Science Programme have changed. Furthermore, these changes have occurred within the context of evolving national research priorities, funding organisations and assessment structure. Beyond both of these dimensions of change, there have been changes in the conduct of scientific research itself. Scientific research communities in numerous other nations now pursue e-Science as the basis for more transformative and accelerated discovery, learning and innovation.

Investments in scientific research, and especially in e-Science, are driven by the recognition that knowledge-based innovation is increasingly critical for economic and social well-being in the globally competitive "flat world."

2. Current Status of the UK e-Science Programme



Introduction

In this section the Panel presents its assessment of the status of the UK e-Science Programme to the present including what it has accomplished in several dimensions: enhancement of academic research, creating a platform (e-infrastructure) for research and economic impact. We also include interwoven commentary on shortcomings and recommendations for the future. These will be recapped in the overall recommendations presented in Section 5. We begin by reviewing some of the complementary objectives for e-Science investment that provide a criteria for reviewing status of the Programme.

The inaugural goals of the e-Science *Core* Programme (eSCP) were to:

- Assist the development of essential, well engineered, generic Grid middleware usable to both e-Scientists and industry;
- Provide necessary infrastructure support for UK e-Science projects;
- Collaborate with the international e-Science and Grid communities;
- Develop a framework in collaboration with scientists computer scientists, and industry to promote the emergence of robust, industrial strength Grid middleware, to not only underpin individual application areas but also to be of relevance to industry and commerce (there was an Open Source/Open Standards requirement on all middleware developed within this Programme).

The complementary goals of the *individual Research Councils*' e-Science investments was to support and leverage the Core Programme investments in service of their disciplinary and interdisciplinary mission of research excellence, impact, public engagement, training of skilled people and provision of world-class research facilities.

An emphasis on *multidisciplinary and collaborative research* with industry, facilitated by e-Science, we infer, is driven by goals to accelerate progress on tackling intrinsically multidisciplinary and complex grand challenge problems. The RCUK definition of such grand challenge areas includes:

- Energy: Secure and sustainable energy supplies are needed to facilitate the economy and have become intrinsic to many people's way of life.
- Living with environmental change: Human activities are accelerating environmental change and increasing pressure on ecosystems and services.
- Global uncertainties: security for all in a changing world (ESRC): Global challenges include the ongoing risks from international terrorism and conflict, rapid technological development, demographic, cultural, and socio-economic change, and increasing pressures on natural resources.
- Ageing: lifelong health and wellbeing (MRC): By 2051, 40 percent of the population will be over 50 and one in four over 65. There are considerable benefits to the UK in having an active and healthy older population with potential economic, social and health gains associated with healthy ageing and reducing dependency in later life.
- Digital economy: Information and communications technology (ICT) has the power to transform the way business operates, the way that Government can deliver and the way science is undertaken to improve life. A World Bank report identified that early adoption of ICT tools supported by research capacity and skilled people better positions a country to reap the economic and social benefits of those tools.
- Nanoscience through engineering to application: This field is maturing rapidly, with a trend towards ever more complex, integrated nanosystems and structures. It is estimated that by 2015, products incorporating nanotechnology will contribute 1 trillion US dollars to the global economy and that the UK has a 10 percent share of the current market.

There is also apparently growing expectation within government that the fruits of publicly-funded research can be more readily and effectively transferred into innovation and economic advantage. There is further expectation that e-Science investments will contribute to this goal.

Phased Development and Status of the Core Programme

The Evidence Document, Section 3, includes a useful overview of the development and progress of the Core Programme in three phases. Below we summarise these phases and the accomplishments in each. The date ranges of the phases are approximate.

Phase One (2001-2004)

- Designated Regional e-Science Centres and linked resources in a Grid.
- Established regional industrial engagement (ongoing).
- Some Centres funded to start training courses and national seminars.
- In 2002 established National e-Science Institute at University of Edinburgh.
- In 2003 a Centre network was established with 7 Centres of Excellence.
- EPSRC funded 6 year CS-oriented interdisciplinary research collaborations. Additional leveraged funding from Core and industry.
- Open call and funding of generic Grid middleware development.
- Established Grid Support Centre and a central Certificate Authority.
- Began international engagement including All Hands Meeting with international participation (ongoing).

Phase Two (2004-2006)

- Created National Grid Service (NGS) in 2004 supported by the Grid Operations Support Centre. Joint funding by Core and JISC.
    - Established single sign-on access to distributed Grid resources.
    - NGS evolved to 4 cores sites, 7 partner sites and 17 affiliate sites.
    - Began participation in European Grid Initiative
- Created Open Middleware Infrastructure Institute (OMII) at Southampton to ensure sustainability of middleware development, quality, ease of use, responsiveness to user needs.
    - Expanded to a network of three by adding Edinburgh and Manchester and linked with OMII-EU.
- Created Digital Curation Centre (DCC) (Core + JISC, now all JISC) to develop best practices and to research key issues in scientific data curation, preservation and open access issues.

Phase Three (2006 - present)

- Beginning 2006, focus on sustaining the activities that have been developed.
- Further funding of the national e-Science Institute (e-SI) at Edinburgh and call for regional e-Science Centres resulting in 7 awards. e-SI shift to longer-term and research-centred themes.
- Established position of e-Science Envoy and appointed Professor Malcolm Atkinson.
- Created e-Science demonstration projects in arts and humanities (new opportunity given formation of AHRC).
- July 2007, e-Science Strategic Advisory Team establishes key focus on:
    - Research into advancing, applying and extending e-Science methods (Human networking grants);
    - Research pioneering novel ICT to enable new e-Science approaches;
    - Investment in infrastructure and support that allows researchers to benefit and exploit e-Science approaches (four Platform Grants + human networking grants).
- Response to above objectives: Four platform grants; four human networking grants; more funding for NGS and OMII. A Call is in process to support management, curation, development of robust software for high quality research.

The Core Programme has supported the creation of important infrastructure resources, training, community building (nationally and internationally) and is now, within the constraints of funding, trying to consolidate gains, enhance take up and use, as well as look over the horizon for more novel technologies and approaches.

Academic Scientific Research

The ultimate mission of the Research Councils includes, perhaps foremost, support of academic research excellence to ensure the delivery of independent, world-class research with impact in a globally competitive world. The Evidence Document includes many details of specific projects within and between disciplines, Research Council funding patterns, and statistical outputs in numbers of projects, leveraged co-funding, bibliometrics, lists of international engagements, number of people supported, summaries of knowledge transfer and collaborative research and eight pages of narrative concerning an impressive array of related funding sources. There is significant affect and support from the research community - at least among the pioneers of e-Science. To supplement and largely confirm the impact of the statistics provided, we now convey a mixture of findings and recommendations of the Panel based largely upon our interaction with the community during the review week.

It is in the academic sector that the most obvious and expected impacts of the e-Science Programme are found. Academic impacts can be divided into three general (although not mutually exclusive) categories: on individual disciplines, on information technology (both as a field of study and as the underpinning of research in individual disciplines) and on collaboration, both within and between disciplines.

The research projects have generated a large number of high-impact journal articles primarily on topics related to the application areas of the projects. Through self-reporting by the PIs with an 84% response rate, since 2001 the e-Science Programmes have supported 966 journal publications, 2054 conference papers, 181 software publications, 103 book chapters and 372 major keynote addresses. The number of citations to the publications in the top five ranked relevant journals are 11,210.

These articles vary in their focus between the e-Science technology aspects of projects and

advancing the state of scholarship in application domains. Researchers and students alike repeatedly told us how influential the e-Science Programme had been on their discipline, offering many individual examples of work that could not have been accomplished; research fronts that would not have been advanced; and partnerships that would not have been formed without these investments by the RCUK. The Panel was presented with a large number of cases providing evidence of disciplinary academic impact.

We also saw much growth in information technology applications within science domains and synergistic activities between computer scientists and e-Science application domains. We did hear some concern about the tension for the information technology e-Scientist who needs to deliver applications that meet the users' expectations and thus the need for development and integration work that does not always produce leading edge information technology research.

Perhaps the largest – and in the long-term most important - academic impact of the Programme is found in the interdisciplinary efforts that have gone into most of the projects and that have fostered new social academic networks and spawned new collaborations. There are excellent examples of how new multidisciplinary groups have been formed around the e-Science projects and the use of the e-Science infrastructures. The long-term impact on the academic communities by this formation of networks is hard to fully quantify, but most certainly should not be underestimated. Evidence of this shift was obvious, for example, in the social-science-oriented projects in which the introduction of e-Science and the emerging interdisciplinary communities have already generated significant impact in terms of access to resources and the creation of new social networks for collaborative research.

The Panel noted that the formation of these new multidisciplinary and distributed communities calls for deep reflection on the practice deployed in the projects, and this forms interesting objects of research for the social sciences. The committee noted the absence, with few exceptions, of projects studying the implications of e-Science on science and the use of social science knowledge within the projects.

It should also be recognised that the interdisciplinary efforts are, generally, harder to publish since the academic community does not yet fully recognise the comprehensive system integration and service provision efforts that go into the e-Science projects to support the strategic application areas. The Panel was provided with several examples of projects in which the information technology participants had difficulties in publishing their results.

In summary, the academic impact of the e-Science Programme must be viewed as very high. The Programme was a forerunner to many other similar academic programmes worldwide, put the UK at the absolute forefront of the development of e-Science and contributed to the acceptance of e-Science as a valid research paradigm. A concern the Panel wishes to express for the continued development of e-Science, both in the UK and internationally, is the lack of academic recognition of the information technology engineering efforts that go into the development of the e-Science software platforms, services and infrastructures needed to support the e-Science application areas.

One example of e-Science Programme impact has been the investments in the social sciences that have been important in two areas: geographic and spatial sciences (GENeSIS) and qualitative analysis (DRESS). The GeoVUE, now GENeSIS, project produced visualisation infrastructure for social scientists and those dealing with the geo-spatial web. The initial usage statistics were impressive. GMapCreator (version 1.1) has been downloaded 9,285 times; PhotoOverlayCreator 1,140 times, and GMapImageCutter 3881 times. However, as the uptake expanded, and GMap Creator became part of the MapTube web site, it has been viewed 27,486 times since April 2008, with the hit rate for the server around 40,000. Similarly, the Credit Crunch Application on MapTube for BBC Radio 4 has had 23,251 responses to the survey. In addition their interactions with BBC1 and BBC2: Britain from Above have had an impact on the British public in areas ranging from interviews to organising and refining vast hordes of information such as the 2001 Census to the intricate network of railways.

Probably the most important impact is the way in which the e-Science investments have situated social scientists to respond to the ESRC grand challenges. Indeed, two of the e-Social Science nodes were the recipients of major ESRC Digital Economy investments, namely in Aberdeen (Policy Grid/Rural Communities) and Nottingham (DRESS/Creative Industries).

Platform for Research Enhancement

The Panel defined the concept of platform to include not just *systems* that meet traditional computing definitions (shared software with interfaces and a variety of extensions features) but also other important forms of infrastructure, including *organisational* (formal and informal groups that provide very important e-Science services), *human capital* (build-up of knowledge and experience that makes it much easier to adapt and adopt technologies), and *data and information* resources (systems that support the increasingly data-intensive nature of research).

Software

An explicit goal of the e-Science Programme was the creation of shared tools. There have been some very important successes. Some tools are in wide voluntary use by researchers around the UK and around the world, and some are moving well in that direction. There is also strong recognition of the difference between a prototype and long-term engineered software that is ultimately required to be true infrastructure.

OMII (Open Middleware Infrastructure Institute) is a serious effort to provide professional support for reused software. The staffing includes around 10 full-time software engineers at Manchester, 7 at Edinburgh and 6 at Southampton. This is a model of professional software development and maintenance that needs to expand. The OMII is a unique service organisation with global importance and impact. It is, however, important to recognise that most application projects we reviewed needed to do significant additional work to tailor the available software to specific applications. This is a consequence of early adopters of the e-Science, especially the Grid infrastructure, trying to use software that had not yet matured to production-level quality. The situation has improved over the years and the efforts taking place within the UK e-Science Programme have been instrumental in reaching a level of maturity which enables more widespread adoption of the use of e-Science infrastructures.

A prime example of a successful software tool resulting from e-Science is the Taverna workflow tool, which has had 65,000 downloads and has been used by 350 organisations. Of note is that Taverna has been used as the platform of a second-level workflow projects for SYSMO. There is now support being provided by a small company. Another important software toolkit is OGSA-DAI (16,000 downloads, used by more than 55 projects).

In a different category, e-Science funded significant work on semantics and ontologies, and supported both research and extensive standards work that had major influence on the OWL and SKOS languages and RDF (all three are World Wide Web Consortium global standards).

Going forward to build on such accomplishments in the initial e-Science Programme, we suggest establishing a process for determining the most important requirements for UK science – distinguishing between needs to support UK-created software and UK-needed software. A classic peer review model of projects is NOT appropriate.

It takes a great deal more effort to support easily usable, reliable systems than to demonstrate clever new functionality (the estimate of a factor of 10 in the classic *The Mythical Man Month* by Fred Brooks is a reasonable guideline). We suggest creating and sustaining long-term centres for software development and support. They need consistent funding at a significant level, not just intermittent support from research projects or the admirable, but inadequate, £1M contribution from EPSRC on behalf of the other science areas.

Software development is not basic research, but instead requires a critical mass of full-time engineering professionals, paid market wages and supported along a genuine career path. There needs to be a commitment to long-term funding and a recognition that software development and support is at least as important to modern science as massive accelerators and telescopes.

There was an excellent summarising comment from the University of Reading (page 115 of the Evidence Document):

Although the Programme succeeded in bringing together computer science groups with other research groups, the collaboration between the two groups led sometimes to misunderstandings: scientists often looked to computer scientists to perform the jobs of software engineers and develop robust tools to order; conversely, computer scientists often - quite reasonably - wish to focus on the latest techniques as part of their own research agenda. The "software engineering gap" continues to be a problem in research computing - who will engineer and maintain the necessary tools? This is an expensive task, and engineering work can be seen as taking money away from research.

In addition, we recommend expanding the technical portfolio in a number of directions that were not as salient a decade ago. These include virtual machines, cloud computing and data-intensive computing. We also advise even stronger international interactions, including consolidation and sharing, especially with EU and US projects.

Hardware

There has been significant growth of hardware investments, through JISC and EPSRC due to the existence and persistence of the e-Science Programme. Networking is a solid area of excellence. JANET and now SuperJANET led the world in richness and availability. There is a risk however that it will no longer be at the leading edge nor sufficient for needs as storage and computational capabilities increase by large factors. Work on wave division and use of dark fibre is important. Now most campuses only have 1-2Gb/s, and the fastest links are 10 Gb/s. Wider connections to the rest of the world also need to be established.

Regional HPC centres and university clusters are connected by various grids (NGSP, GridPP, UK-QCD) to provide a good model of a multi-tier capability. At present the highest performance machine in the Grid is the national HPC Centre's HeCtor which is a 200 TeraOp Cray XT5 with 1PB of storage. This is no longer an impressive top-tier national resource and the storage is much too small.

We strongly urge a top down analysis of what are now the most cost-effective means of providing basic storage and computation. Specialised hardware is still needed for some sorts of calculations, but many sorts of data-intensive and massively parallel computation can be done on modified commodity (i.e. less expensive) hardware. There can be significant savings based on scale, and it may be time to revisit having more centralised national scale storage and computation resources, with greater bandwidth connections from the user community than now planned. Latency will not be a problem for most uses within the UK. The skills needed to run such data centres are quite different from those for developing software and technologies; the latter can be done at a distance.

The HPC and Grid infrastructure are important vehicles for e-Science research, but other architectures can also play this role in the future. It may sound trendy, but consider the potential effectiveness of cloud-like services to provide modest (1-10 gigabytes) of guaranteed storage for individual researchers and small projects, with terabyte- and petabyte-sized managed storage for larger projects (on an allocation basis), along with application and compute on-demand services.

Human Capital

A key positive outcome of the e-Science Programme has been that many people have become experienced with computational thinking and technologies and experience in cross-disciplinary research. We saw considerable evidence of the results of investments in people. Some examples include:

- Aberdeen said they bid on and won of the three very large national Digital Economy Centre grants and that they would not have been able to do that without their e-Science experience and collaboration across disciplines on campus (all three of the digital economy awards, £12M each, went to sites that had had e-Science centres).

- A quote from a medievalist at University of Sheffield: "we could not have done this work without the e-Science Programme and the e-Science funding" (new platforms for analysing manuscripts, new ways to establish authorship, etc.)

- In Oxford, the Panel saw the results of palaeographers working with medical imaging experts, without a traditional computer scientist in the mix.

- In York, Jim Austin said that "the DAME project would not have happened without e-Science funding."

Besides examples of learning by individual researchers, we also saw examples of growth of capability in larger groups:

- The Oxford VOTES project has learned how to run massive clinical trials efficiently, moving from a cottage industry to an automated factory model with specialised software and hardware. As a result they have been paid to run trials outside of the UK based on that expertise.

- The GridPP group has developed considerable expertise in the methods and technology required for large-scale, distributed data management and analysis.

Furthermore, we saw many opportunities for training, regional development and the encouragement of new scientific disciplines. e-Science funding has been very important in training students in new technologies, sparking new MSc-level programmes, and new interdisciplinary Doctoral Training Centres. We heard definite evidence that e-Science had trained students in ways that are quite valuable to industry and that they have been easily employed. Unfortunately, there does not appear to be sufficient hard data on careers for the Panel to make quantitative statements; we recommend collecting better data on alumni.

Examples of industrial and regional impact include:

- Many Newcastle graduates are employed in local industry, where an absence of skills was holding back development.

- Rolls-Royce/OSyS started a project using research in pattern matching and measurement and data management, extended it to Grid Usage, and is now moving to the BROADEN project with wide

Rolls-Royce intraGrid and even corporate extraGrid use.

- First Derivatives in Belfast has adopted grid technologies with help from BeSC.
- The BBC established a new division based on their work with BeSC.

The continued development of e-Science methods and their application within UK science must ultimately depend on the next generation of researchers. In this regard we wish to emphasise the critical importance of ensuring a viable career path for young researchers working in e-Science. We find that this career path is on shaky ground at present. There are the traditional difficulties about recognition for cross-disciplinary research. This sort of work is, however, crucial for the vibrant growth of science, since many of the exciting results come at the interfaces of existing fields.

It seems that the e-Science Programme allowed a few young people to progress from PhD student onto the lecturer track and hence to more senior positions. However, the decline in the e-Science Programme means that people who entered later find themselves employed on uncertain RA funding. This situation is problematic for two reasons. First, the current generation of talented scholars will be lost to the UK research enterprise. Furthermore, people with strong analytic and software skills are in demand worldwide, both in academia and industry (as more than one person observed, when asked if these people would continue in e-Science, "Yes, but not in the UK"). Second, it will be hard to motivate new students to enter the field if there is no ongoing upward path for them.

We appreciate that it will never be possible for more than a few PhD students to obtain academic posts. However, there is a particular need to ensure that some posts are accessible to e-Researchers. The Research Councils can encourage this by ensuring that a good number of their Fellowships (Career Acceleration Fellowships, Leadership Fellowships, etc.) are dedicated to interdisciplinary researchers. We recommend that 20% of all fellowships across the councils be dedicated to e-Researchers. We also urge collection of information on the (mostly non-academic) careers of graduates of e-Science master's and doctoral programmes.

Community, Organisations, Institutions

The e-Science Programme has driven and supported greater community building and institution creation that have significantly addressed the e-Science goals. We gathered evidence of significant creation of formal institutions and organisations as well as informal communities. Some of the e-Science institutes are continuing, even without ring-fenced funding, albeit at a modest scale. There are Doctoral Training Centres specialising in e-Science. Some universities have created their own organisations (e.g., the Oxford eResearch Centre). As a result of the e-Science Programme, new fields and academic specialties have sprung up. For example, York now has a degree programme in Computer Science in Business Enterprise Systems.

The annual e-Science All Hands Meeting continues. It has been a key social mechanism and a way to bring potential partners together. Although it is no longer financed by a central e-Science organisation, it continues to run and attract over 350 attendees. Some of the e-Science centres continue. The National e-Science Centre, joint between Edinburgh and Glasgow runs many workshops, tutorials and other events, with thousands of visitor-days each year.

Data and Information Stewardship

Another important infrastructure impact is the increasing importance of data and the way that e-Science promotes treatment of data as a resource and the emphasis on the collection, curation and processing of data. The data repositories created in e-Science projects will continue to have an impact over many decades. The build-up of an infrastructure capable for handling the current and future demands of the data-intensive sciences is, without question, one of the most important long-term impacts the Programme has generated. However, we are concerned that inadequate investment has been made in the actual curation of the data produced by e-Science projects. The Digital Curation Centre serves as a convening entity to provide expertise, acts as a clearinghouse for best practices in data management, and sponsors an important conference and a peer-reviewed journal. These are essential investments, but they are insufficient to provide the long term access to the research data of e-Science projects, such that their value can be leveraged by mining and combining extant data.

The e-Science Programme recognised early the key role of improved access to information and supported relevant research and tool building. There have been specific investments in storage management at the discipline level, but we saw no plans for large-scale support of storage or federation of data resources.

There has been major progress in the understanding and provision of metadata. Explicit schemas and metadata definitions are now common in many areas. One example was the radio telescope project which budgets 10% of their massive data as metadata. We have already commented on the important research and standards work on ontologies and semantics.

The issues of storage and information management for the sciences are very difficult and involve both research and policy decisions. These have not been solved outside the UK yet either. There is an opportunity for leadership and further research as well as clear-sighted policy and finance decisions.

There must be broad planning for what should be saved and how it will be made accessible and usable over time. To enable future cross-disciplinary research, the Programme needs to ensure pieces are saved in the disciplines, even though they may be considered of limited future use to the discipline. There is also a need for an information and knowledge strategy too, including linked ontologies, standards and metadata stability.

Most important, there must be a very serious look at the expectations for fast, exponential growth of data that are created and the value of the information to be conserved. There is an inevitable conflict between saving results and doing new work, so there should be overt and transparent decisions. Such decisions involve significant capital and operating expenses, though many of these are hidden at the local level. In addition, decisions about what to save are sometimes made locally and based, not on the long term value of preservation, but rather on immediate storage shortages.

Services

The UK e-Science Programme, in support of modern e-Science needs, has created and operates some heavily used services. Some important continuing services are:

- Network operations: As commented above, SuperJANET has been a major success and is essential to the entire collaborative and distributed approach to computing and to research. There are needs for operation of the facilities as well as paying for them and planning their growth.

- The National Grid Service (NGS) has 25 member institutes, 33 resources, 15K processing cores; in the last year the NGS provided 4.7M processor hours to 900K jobs. A considerable fraction of NGS usage is by biologists (BBSRC) – something not seen when the first Grids were provisioned. Other grids include GridPP, UK-QCD, White Rose Grid, ScotGrid and NERC data Grid. "The goal of the NGS is to deliver a production quality national e-infrastructure in support of academic research across all Higher Education Institutes (HEIs) in the UK."

- JISC provides authentication and security services which are essential to reliable distributed computing. It has issued 22,000 certificates (almost 5,000 are currently active), and it supports the National Shibboleth federation and the SARoNGS service for Shibboleth-controlled access to NGS resources.

- In support of collaboration, the Access Grid continues to operate, with 1200 events per month. More recently a video-conferencing centre was established. These have enabled far-flung projects to keep coordinated and advance their work.

- Digital Curation Centre (DCC) does laudable work on end-to-end lifecycle analysis, consulting and advising, and provides essential contact and representation roles to other organisations. There is not yet a national curation service and there is opportunity for considerably more direct engagement with science projects to provide operational services. A recent *UK Research Data Service* report along these lines is available at **http://www.ukrds.ac.uk/**.

It is important to sustain essential operations for use by all. Unfortunately, currently there is no plan for long-term basic support outside of R&D project budgets. Researchers must be able to count on services continuing and providing a reasonable level of quality. We advise starting an effort to revisit the decisions on which categories and specific services need to be provided specially or can be procured commercially, and then make a strategic plan working together with the research community, service industry and government.

We suggest addressing the most difficult aspects of information services (see the data section above). We also advise a hard look at the collaboration services. In the current state of interactive systems, many basic services may be available as commercial or commodity offerings, but there are new high-end possibilities that might also be investigated.

More on the Challenge of Sustaining and Improving the Platform

The UK e-Science Programme as a whole has yet to create the right environment for sustaining and continuous improvement of important e-Science research facilities that have extra-ordinary levels of utilisation by their respective communities and high status within their communities. There were a significant number of exemplary services that have been created during the e-Science funded period that have revolutionised the delivery of these services and have been adopted by the national and international community. Example services are in genomics, taxonomy, oceanographic and atmospheric services. These services often brought together data sets into collections that clearly are leading to research and research discoveries in key government priority areas, but in themselves are unlikely to have research outcomes. As a consequence these services are not considered appropriate for funding through normal competitive research funding processes.

While these services serve their respective disciplines, they are not easy to fund from the traditional Research Council funding arrangements. Experience shows that these research services often have significant impact outside their disciplinary area because they provide products that can easily be used outside the discipline and therefore they support multi- and cross-discipline research. These services lead to greater reuse of data and knowledge, and therefore magnify the value of the original data, and enable new research questions to be explored.

While sustained funding of research services is a weakness in many countries, the UK is in a unique position, where some Research Councils have recognised the value of providing facilities that support the research community (e.g. British Atmospheric Data Centre or British Oceanographic Data Centre funded by NERC). Other services were shown which do not have support from their respective councils, and it was clear that sustaining these exemplary services was at a crisis point. Some countries have already created national services to support research infrastructure. For example, Australia has created a programme specifically to create research facilities that support researchers within and across communities through the National Collaboration Research Infrastructure Scheme (NCRIS).

Economic Impact

Quantifying the economic impact of the £214-plus million investment in e-Science faces the same set of challenges as any other large-scale and complex scientific investment. These include: capturing the direct impact of expenditures in the science sector; describing the multiple ways in which scientific investments directly affect economic activity through the transfer of technology; and delineating the multiple indirect ways in which the economy gains through scientific knowledge, such as better health, reduced vulnerability to climate change and better communications. The challenges are exacerbated in the case of infrastructure investments, which may have only indirect and substantially lagged effects. Finally, the most difficult challenge is identifying the opportunity cost of the investment that is necessary to determine an appropriate return on investment.

In the case of the e-Science Programme, the challenge is more difficult in that grant recipients were not asked to keep track of their impact from the inception of the Programme. Thus the discussion below should be seen more as a qualitative description than a formal analysis of the impact.

The Panel concurred that the UK investments had demonstrated that innovation, design, resource management and decision support could be revolutionised in the full design-to-delivery product life cycle. The impacts are already being felt in three important areas: life sciences and medicine, materials, and energy and sustainability.

The Panel identified three canonical ways in which projects had a substantial economic impact, that may help inform the design of future investments. The first of these was direct involvement with industry from the inception of the project, like the DAME/Rolls Royce project, where the driving research question was in response to industry needs. The second was aggressive, university-led linkage with industry during and after the development of new technologies, exemplified by the Imperial College and Newcastle approaches. The third was due to extraordinarily entrepreneurial PIs, such as Carole Goble at the University of Manchester, and Tom Rodden and Andy Crabtree at Nottingham University, who identified and exploited market opportunities. Newcastle's Digital Economy Hub is also founded on their e-Science work led by Paul Watson.

Direct Economic Impact on the Science Community

The direct impact on jobs and employment is documented in Section 3 of the Evidence Document. Between 2001 and 2008 there were at least 162 postgraduate students employed, and 1,516 academic staff funded through the Programme. Although the academic production function is such that any infusion of resources would be expected to generate roughly this amount of faculty and post-graduate support, the Panel ascertained that the additional value associated with the interdisciplinary training of graduate students was an important economic outcome that would otherwise not have taken place. For example, students trained in e-Science at Nottingham University achieved placements at Microsoft Labs in Cambridge with salaries about 15% higher than would have otherwise been the case.

Direct Impact on Industry

The attractiveness of e-Science to industry was evident from the inception of the Programme: it attracted £20 million in industrial collaboration and £7.1 million in cash and in-kind industry transfers. The initial results are quite promising: the Programme has resulted in 138 stakeholder collaborations, 30 licenses or patents, 14 spin-off companies and 103 key results taken up by industry.

The Panel identified a number of examples in which the Programme had led to substantial advances in regional, national/European and international economic competitiveness.

In the case of Newcastle, for example, the e-Science Programme led to substantial regional impact through collaboration with Arjuna Technologies. Seven PhDs and thirty MScs from Newcastle University have worked for Arjuna, one PhD, many MScs and a Business Development Manager have returned from Arjuna to Newcastle University, and the institutionalised network of knowledge transfer to local companies has led to "cutting-edge technologies:" e.g. web services, grids and clouds and the regional involvement of Red Hat, Amazon and Microsoft. Newcastle University estimates that Arjuna/RedHat have contributed ~£16M GVA to the regional economy.

Nationally, the Distributed Aircraft Maintenance Environment Project (DAME) partnered with Rolls-Royce, Data Systems and Solutions and Cybula Ltd to use e-Science to reduce engine maintenance times and to improve the interoperation of the maintenance team. The technologies developed are now used on Rolls-Royce Trent engines and the result was a spin-off company: Oxford Bio-Signals (OBS).

At the national/European level, Imperial College's DiscoveryNet project led to the formation of a company that was acquired by IDBS to form the largest European e-Science company for scientific information management and analysis with 250 employees and annual revenue exceeding £20M; InforSense was formed as a successful spinout to commercialise Discovery Net technology. The customer base is now over 100 companies, with £25m total revenue in 8 years. It is both the key technical provider for 6 EU FP6 projects, and ranked by top industry analyst firms such as Gartner as the key vendor for advanced healthcare IT infrastructure. Other e-Science investments in Imperial College also led to partnerships with Syngenta, O2, Vodaphone, GSK, Transport for London and the UK Department for Transport. Similarly, the investment in GridPP led to the establishment of the start-up companies' imense and iLexIR at Cambridge (Camtology). GridPP/ATLAS/LHCb developed the Ganga interface, as well as Econophysica (mathematical models for commodity trading) at QMUL, with access to resources at Total Oil (geoscience research) from Aberdeen.

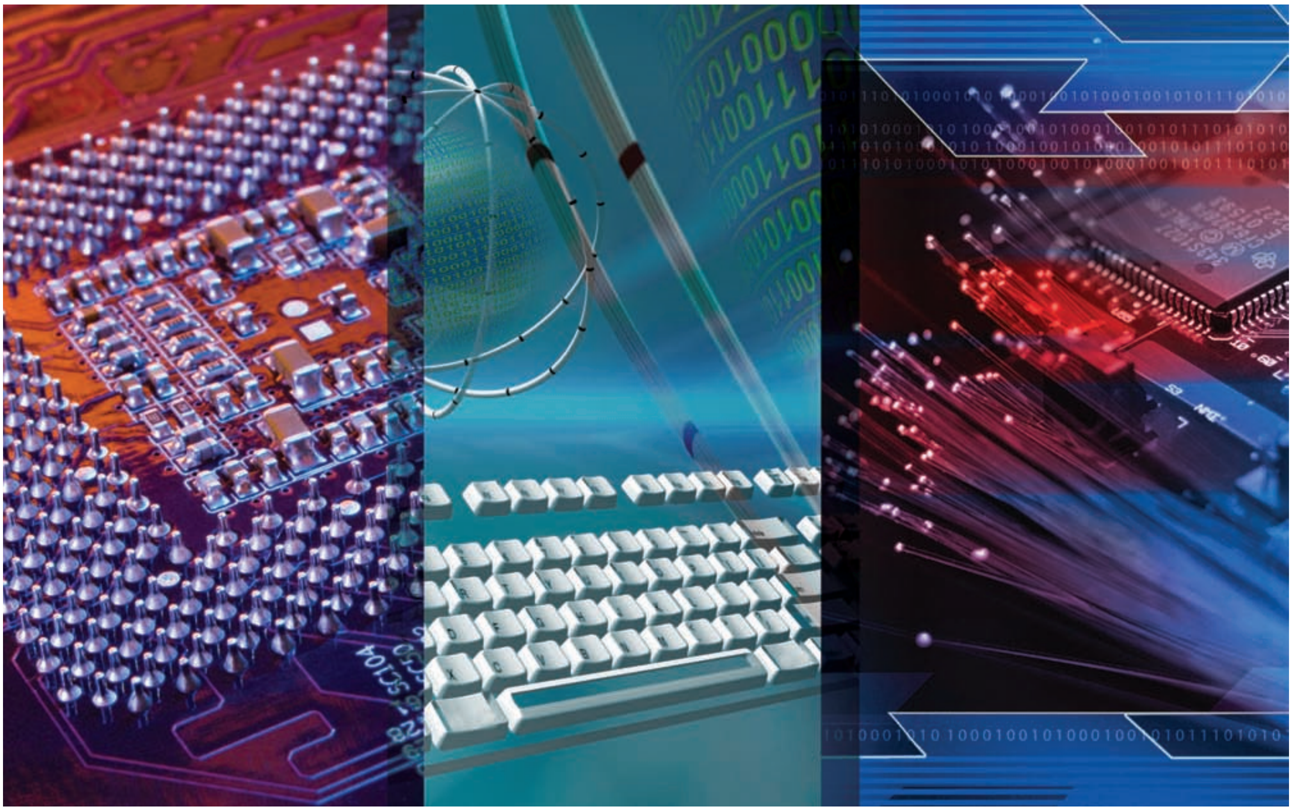
At the international level, the University of Manchester's development of workflow systems has been adopted by more than 350 organisations (and 23 companies) in 35 countries in the UK, China, Europe, USA, Canada, SE Asia and South America. In the five year period between July 2005 and June 2009 the whole consortium acquired £13,144,651 of funding: £4,990,316 EPSRC; £6,316,693 other funding councils; £1,057,817 from the EU; £600,839 from the USA; and £715,839 from industry. A strong commercial usage by small and medium sized enterprises and private institutes in drug discovery and therapeutics services has lead to a link with a commercial company, Eagle Genomics, to provision commercial support services. Links to another commercial company, Informatics-Clinical Information Systems, has led to new products for the healthcare sector.

The Belfast e-Science Centre had industrial impact with at least three projects: Gridcast with the British Broadcasting Corporation (BBC) and British Telecom (BT); the UK's first financial services grid project OpenRiskGrid with First Derivatives; and Grid Enabled Distributed Data Mining (GEDDM), for the security domain which enabled new commercial services to be deployed by Datactics Ltd.

Secondary Effects

Almost all the projects that were studied had the potential for important secondary economic impacts on health and social wellbeing, although in many cases, it is too early to identify the full effect. The investments in GridPP at RAL and the Environmental Systems Science Centre at the University of Reading provide illustrative examples. There is likely to be a substantial secondary benefit from the avoiding the impact of lost work days as a result of the ability to rapidly screen many drugs "in-silico" in the way that GridPP resources were used to screen potential agents in the fight against bird-flu and malaria. Similarly, the University of Reading's research to improve the prediction of extratropical storms by numerical weather prediction using storm identification and tracking software led to a reduction of the flood and wind related damage due to storms.

3. Future Considerations: Vision, Opportunities, Risks, Timescales and Response



Introduction

When the e-Science Programme started, major bets were made on a number of risky but promising technologies. These included:

- Superior networking (SuperJANET)
- Distributed computing
- GRID (technology for secure federation)
- Data curation, sharing and management
- Semantics
- Process and workflow management
- Web technology

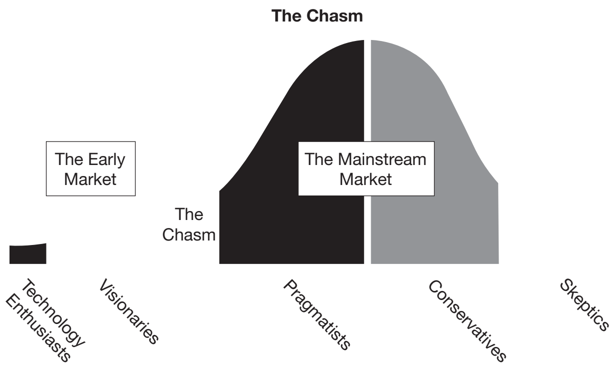
Considerable progress has been made on each of these. Research has been done, experience has been gained, tools have been built, infrastructure investments have been made and e-Science projects have been run using them. In most cases these have turned out to be excellent choices showing great forethought. It is reasonable to reconsider these, 6-8 years after the original decisions, to decide which deserve further investment and support and what other directions may be worthy of being started. But each of these has prospered sufficiently that a plan for sustainment as well as an approach to broadening use is essential.

These technologies are already part of the research modus operandi for many scholars, and they have increased productivity and capability enormously. They promise even greater value as they spread throughout the community – but only if there is a credible plan for long-term support so that researchers can count on their availability. This requires consideration of technologies that are both created and used by UK scientists. The general market outside e-Science will not be adopting most of these technological resources sufficiently quickly or broadly to make them easy off-the-shelf purchases; the needs of the scientific and scholarly communities are different enough, and in many cases extremely so, that they demand direct approaches.

It is also important to realise there is a time dimension – different technologies are at different states of technical readiness and in different stages of adoption.

Crossing the Chasm

A useful model for understanding the situation comes, strangely enough, from the high-tech marketing literature. Geoffrey Moore's 1991 business classic, *Crossing the Chasm*, uses the standard description of the Technology Adoption Life-Cycle.



"... the point of greatest peril in the development of a high-tech market lies in making the transition from an early market dominated by a few visionary customers to a mainstream market dominated by a large block of customers who are predominantly pragmatists in orientation ...

The *early adopters* buy into a new product concept very early in its life-cycle ... they are people who find it easy to imagine, understand and appreciate the benefits of a new technology and to relate these potential benefits to their other concerns.

The *early majority* share some of the early adopter's ability to relate to technology, but ultimately they are driven by a strong sense of practicality...

The *late majority* shares all the concerns of the early majority, plus one major additional one: Whereas people in the early majority are comfortable with their ability to handle a technology product ... members of the late majority are not" (Moore, *Crossing the Chasm*, 2nd edition, pp 11-12).

Recasting these ideas for the research community: The visionaries are willing to adopt pieces of technology that will allow them to do things they never could before, despite the risks and difficulty. The pragmatists on the other hand want a reasonably complete package of capabilities that they can use to solve problems better, but they do not want to spend much time or effort folding it into their work.

The hardest part is moving beyond the original set of enthusiasts and visionaries, who see the excitement of the new and its potential to allow them to solve problems in new and exciting ways, to people who simply want to take advantage of (perhaps radically) improved technologies to do their research, but who do not want to divert their attention unduly to the new tools rather than their own domain interests.

The great risk today: If the technologies do not make it into the mainstream, they will not realise the early promise and expectations nor pay back the investments in research, staff, money and hope. It is essential to make the organisational and financial investments to provide the support for necessary technologies and to ensure they reach the mainstream of research as technologies and processes of choice.

Sustaining Requirements and Resources

Superior networking

An area where the UK has a long-term success plan is networking for science and research. SuperJANET provides considerable bandwidth (1Gb/s and 10Gb/s links) among the major centres. There is also a planning and provisioning organisation and a technology testing group exploring a model for the future. We do urge re-examining the intensity (number and speed) of practical links as requirements increase exponentially for remote computing, visualisation and especially data storage.

Distributed computing

Distributed and remote computing is a reality for the UK research community. It links shared resources among UK researchers and provides access to computational, data, and observatory resources in the US and the rest of Europe. There needs to be continued investment to ensure resources are shared and scheduled responsively. There needs to be renewed examination of the distributed storage opportunities, including massive storage clusters (many petabytes at minimum) that would support preservation and long-term access to data and information.

GRID architecture

Distributed and remote computing is a reality for the UK research community. They share resources among UK researchers, access resources in the US and the rest of Europe, and provide such access to others. The UK adoption of Grid technology (e.g., Globus Toolkit) has facilitated these modes of use and the UK has also contributed to the maturation of this technology. The UK has played a major role in standards setting for GRIDS being highly active in GGF (and then OGF). The UK chaired many of the research and working groups, like SEM-GRID and OGSA-DAIS, and were instrumental in standards like JSDL and SAGA.

UK efforts on federation and sharing of compute resources have been focused around regional grids, the National Grid Service and the GridPP project. These efforts have been highly successful, providing many users with access to more computing power than they could otherwise easily obtain. Looking forward, we recommend that these efforts, including enhanced capacity and function of distributed storage, be sustained and expanded.

UK efforts aimed at furthering the publication of data and software resources as services ("service-oriented science") have been less successful. Early experiments with the use of Web Services technologies for exposing data and software as services (e.g. via GT3) met with difficulties due to immature technology. The technology has since matured (in GT4 and beyond) but large-scale examples of service oriented science remain few in the UK: there is nothing on the scale of the US cancer biomedical informatics grid (caBIG), for example (projects at the Belfast e-Science Centre and perhaps a few others are examples of exceptions). We recommend that the UK e-Science Programme look carefully at how to incentivise broader participation in data and software publication and to join in collaborative projects with international partners aimed at bringing the most modern technologies to bear on this task.

Data curation, sharing and management

An excellent step was taken in setting up the Data Curation Centre which provides consultative services and represents the UK in international standards and other contexts. Curation is about management and about meta-data as well as resources. This is still an area of spotty impact and one that may call for a different level of investment to move the scholarly community to a model of broadly available information that can be accessed over long periods of time and among many disciplines. This may require an order of magnitude change in expenditure and effort to cross the chasm. Finally, although a lot of progress has been made, security still has to be addressed in the context of sharing data and e-Infrastructure, particularly the relationship between commercial security strategies and the research communities current practice.

Storage management

Implicit in the data curation mandate is support for the storage of the information. The commercial arena has made great strides in managing very large amounts of data and doing it economically. It may be appropriate to examine these approaches in the context of the data curation needs.

Semantic technology

Making formal semantics a practical technology was a daring idea that was embraced by the e-Science Programme. There has been considerable progress on research (and the UK has been at the leading edge) and on creation of standards. These have provided a base for rigorous ontologies in a variety of fields. Semantic technologies also are a support for the successful workflow engines that have come from e-Science. Decisions need to be made on how much further to push the technology development and what needs to be done to cross the broad chasm.

Process and workflow management

The e-Science Programme has created viable tools for handling moderately complex and voluminous workflows. These tools have spread fairly widely and have been used by groups remote from the original researchers. This area may be mature enough for long term-support.

Web technology

e-Science Programmes make appropriately heavy use of the Web. The abilities of the Web have of course changed radically over the period (Web 2.0, social networking, etc.) so it may be time to rethink what aspects of Web technology are most in need of evangelism, development or specialised support for scholars.

Investments and Adoption

Crossing the chasm requires a different sort of support and drive than staying with the visionaries. There must be a full set of software, systems, processes and training to enable ordinary users to use the new approach. There need to be people willing to hold hands, commiserate or share their own experiences. There must be a reasonable promise of ongoing support and availability of the capability. Staff who

understand the technology (often researchers in their own right) can talk to scientists and solve problems – and who also have the right mindset, incentives and career paths – must be available in an ongoing way. These people may be expensive and require management and long-term organisational support that is atypical for universities. The e-Science Programme has supported several middleware centres with these characteristics. We are urging larger and more permanent organisations which may not be part of university research departments. Some of these needs are well covered, frequently under the joint infrastructure umbrella of the RCUK, but others are not, and in the current state utilise patched-together staff and time from the research and early user groups.

It is important to remember that there are two sides to the investment and adoption issue: (a) what UK e-scientists need, and (b) what UK ICT people produce. There needs to be cooperation and collateral learning between the two. There also needs to be a willingness to use good middleware and application codes produced outside the country and to avoid wasteful not-invented-here attitudes. The UK Programme has made heavy use of Condor, Globus, SRB and other foreign tools, acquired at little or no direct cost, but at the same time put considerable effort into developing indigenous technologies, sometimes in competition with foreign tools.

The Next 5 Years: From Research to Sustainability

To provide more specific vision and guidance for the future, based upon the information provided to the Panel, we present a 5-year forecast and a 10-year vision of what the e-Science Programme could look like in the UK.

Maintain critical-size centres already established

The Programme has been spectacularly successful in establishing several critical-size centres (Oxford, Glasgow, Edinburgh, Newcastle, Cambridge and others). The key to the success of this part of the Programme was the strong partnership between the RCUK and the individual Universities. Clearly an investment of the order of £5M has generated considerable leverage: establishing new programmes, opening new tenured academic appointments and establishing critical mass at the large centres. It is the Panel's judgement that if we considered the top-ranked e-Science institutions around the world, the top would be dominated by the UK. However, unless there is a stable, long-term continuation of such targeted support, there is a danger that the current efforts will be diluted and eroded over the next 5 years.

Dramatically expand community involvement

Community buy-in is clearly emerging. The current Programme has attracted the early adopters and has proved the success of the e-Science approach. A continuing investment at least at the current level over the next 5 years could easily grow the involvement by at least a factor of 10. This should happen in the research community, in linked industry and in society as a whole. Many of the e-Science skills (data analysis, data-intensive statistics, data management, simulation, modelling, collaboration) are very much needed by a wide range of governmental institutions and commerce alike.

Provide mid-term career paths for current personnel

The current Programme has been very successful in engaging young scientists in e-Science. Many of them graduated but relatively few of them were senior enough to make it into a permanent position. From our interviews it was obvious that there is much angst about their future. Much of the current investment will lose its long-term impact if the whole generation were to leave the field. There has to be some dedicated funding to ensure continuity in the career path of this group.

Build stronger bridges to the high performance computing (HPC) community

The e-Science Programme was deliberately designed to be clearly distinct from the more established and traditional High Performance Computing. We believe that this was an important factor in the success of the current Programme, enabling adequate independence so that it could build up its own character. However, now bridges need to be built and mutually beneficial projects implemented. Good examples are the efforts in Oxford and RAL.

Develop shared infrastructure that is reliable, mature and sustainable

Traditional trends are turning upside down. Hardware (the traditional capital investment) is becoming obsolete every 2-3 years, while software is becoming the capital investment. The standard scientific practice of constantly reinventing and rebuilding existing tools is becoming increasingly untenable. The best way forward is to develop a shared infrastructure that is of high enough quality to be used across multiple disciplines and is designed with modern software methodologies so that it is sustainable. The National Grid Service (NGS) has done quite well in establishing the basis of such a maintainable effort. This is rather unique to the UK – no other country has started a crosscutting effort to establish such a shared infrastructure – and deserves to be nurtured.

The Next 5 Years: Sustain and Grow the Emerging Community

Strong national leadership, stable cross-council coordination

The current effort would not have been possible without the strong leadership and distinct vision of Sir John Taylor, Professor Tony Hey and Professor Anne Trefethen. A group of very competent PIs have seized the opportunity these people helped create. The result has been spectacular, but it is increasingly clear that simply cutting off the Programme will not yield the expected result. It will take a much longer period to grow the whole effort to be an organic part of UK science. The current success of starting such a project speaks for itself; it would not have been possible without the initial strong cross-council coordination. e-Science is special – it is not one of many crosscutting projects but rather is building the fundamental instrumentation of the 21st century that will be common to most scientific disciplines. It is also a platform that will ultimately enhance most all knowledge-based activities – both discovery and learning.

Community building and training

The current project has established some very effective organisations, the National and Regional e-Science Centres and the yearly All Hands Meetings. These need to be continued, evolved and broadened. Furthermore, efforts to create Doctoral Training Centres, developing systematic curricula for e-Science, must be carefully fostered. Having summer schools for doctoral students might be a very efficient way to build up a community.

Systematic dissemination of best practice

There has to be a dedicated knowledge transfer organisation that deals with disseminating "how-to" information within the e-Science community and would provide help in jump-starting new science projects. Grants should be rewarded, not punished, for reusing existing components.

Packaged and hosted service-based e-Science

Some of the best e-Science tools and practices can be made available through simple web-based services, which would enable a very low threshold for newcomers. These services can be various statistical tools, advanced visualisations, data federation services (AstroGrid, OGSA-DAI), data curation tools and metadata/semantic services.

The Next 10 Years: Evolutionary Challenges

Keep up with emerging and evolving new technologies

Data intensive computational technologies are evolving extremely rapidly. A decade ago we saw the transition from supercomputers to BeoWulf clusters and then the Computational Grid emerged. Today we are exploring Cloud Computing, GPU based computing and data-intensive computing. The UK effort will need to be "running forward" just to hold on to its leading position in the world. These emerging new technologies need to be constantly evaluated and tested on pilot projects. Wide deployment of relatively inexpensive sensors, tied to computational tools, have been responsible for many of the recent scientific breakthroughs (gene-chips, sequencing, remote sensing, sky surveys, medical imaging, etc.). Setting up a shared, national support centre steadily fulfilling these exploratory roles would be a very cost-efficient way to approach this unavoidable challenge of future developments in e-Science.

Transformations in different disciplines will not happen at the same time

There is a continuum in how different scientific disciplines have adopted the e-Science paradigm. High Energy Physics and Astronomy are quite far advanced in embracing e-Science while other areas need more time. It is not unrealistic to expect, given academic turnaround timescales, that over a 10-year period most disciplines will have been exposed to e-Science in a much more systematic fashion and e-Science will become as deeply ingrained in the sciences (and arts) as originally hoped for.

Stable hierarchical distribution of resources

The current e-Science Programme has been very successful in planting the seeds of a stable and robust system. The UK effort has built up a hierarchy of centres and resources ranging from groups to universities, regional and national centres, while also participating in large international efforts. This is totally unique in the world, and enables a much more gradual evolution of the system and support infrastructure than any other national effort we are aware of. This should not only be preserved but systematically grown and evolved over the next ten years, placing new, large scale facilities and resources at the appropriate levels of this hierarchy.

The Next 10 Years: Data Challenges

Data sets: new kinds of instruments

Data sets are becoming the new instruments of science. This is obvious just by looking at the current data holdings and their use at various UK e-Science centres and Institutes. The high-throughput sequencing data at the Wellcome-Sanger Institute is already reaching into petabytes. Federation of oceanographic and astronomy data enables qualitatively new science and the reuse of electronically archived medical information is already revolutionising health care. A thousand years ago science became empirical; in the last few hundred years it evolved a theoretical branch; in the last few decades a computational branch; and we are evolving a data exploration branch. A new, *fourth scientific paradigm* is emerging, centred around innovative uses of scientific data (see **http://research.microsoft.com/en-us/collaboration/fourthparadigm/**).

Self-amplification of data

The easier it is to collect and analyse large amounts of scientific data, the more data will be collected. New areas in social science and the humanities will become increasingly quantitative and use much larger data collections that anyone thought possible. Many new data sets will emerge from the federation and fusion of various data sources, and combined in unique ways, they become new, value added data sets in their own right. Data sets are generating new data sets themselves, like "real instruments" do. Scientific data is currently doubling every year, and over the next 10 years this doubling time will shorten to 3-4 months.

Simulations larger and in more disciplines

Computational simulations have been for a long time part of engineering and physical science. Recently biology has entered this field; simulation of complex organs like the heart are done to the utmost detail. Scientists working in systems biology are attempting to simulate the inner workings of a living cell. Social sciences are using increasingly more sophisticated techniques to simulate social phenomena, soon to the level of including individuals in a UK-wide simulation. These tools generate even more data and will cause a substantial pressure on the common infrastructure. At the same time this is the area where there will be a convergence between traditional high performance computing and e-Science.

Data growth beyond any current imagination

It is difficult to project data sizes that will be commonly used in sciences in 10 years from now but clearly they will be in the hundreds of petabytes, beyond anyone's reach today. Dealing with such amounts of data will be very difficult even on the national scale, impossible locally. Having a balanced hierarchical system in the UK will be an enormous strategic advantage compared to other countries and potentially the only way to address these challenges successfully. The challenges from this data-intensive future are also about more than scale. They are also about handling more complex, heterogeneous data and more complex analyses.

Special Opportunities for the UK in e-Science and Social Sciences

Both in the UK and internationally, e-Science tools have not been widely adopted within the social sciences. There is a set of "early adopters" with a particular technical inclination and research interests that are demanding in requirements for data, collaboration or computing. However, social science research is on the verge of being transformed through distributed global collaborations, the use of very large data collections, terascale computing resources and high performance visualisation in a way even more fundamental than research in the physical and life sciences. Simply put, the ability to capture vast amounts of data on human interactions in a manner unimaginable from traditional survey data and related processes should, in the near term, transform social science research. For example, a person's interests and networks can be uncovered through the online behaviour documented by the major search engines, such as Yahoo! and Google, as "data collection events."[2] Geographic movements can be tracked by cell phones which include GPS location information.[3] Health, work and learning information can be tracked using administrative data from hospital records, employment records and education records.[4]

The impact of social science on both economic and social policy could be transformed as a result of new abilities to collect and analyse real-time data in a far more granular fashion than from survey data. In the United States, such technologies are already being used for research purposes to great advantage. For example, Schunn uses video data collected from a recent highly successful case of science and engineering, the Mars Exploration Rover, to study the way in which human interactions contributed to the success of the project. While the project both wildly exceeded engineering requirements for the mission and produced many important scientific discoveries, not all days of the mission were equally successful. Schunn uses the video records to trace the path from the structure of different subgroups (such as having formal roles and diversity of knowledge in the subgroups) to the occurrence of different social processes (such as task conflict, breadth of participation, communication norms and shared mental models) to the occurrence of different cognitive processes (such as analogy, information search and evaluation) and finally to outcomes (such as new methods for rover control and new hypotheses regarding the nature of Mars) (Schunn, 2008).

Human behaviour is increasingly captured through transactions on the Internet. For example, most businesses, as well as registering with the tax authority, also create a website. It is now entirely

[2] http://bits.blogs.nytimes.com/2008/03/09/how-do-they-track-you-let-us-count-the-ways/?scp=17&sq=privacy%20yahoo!&st=cse accessed Sept 19, 2008.

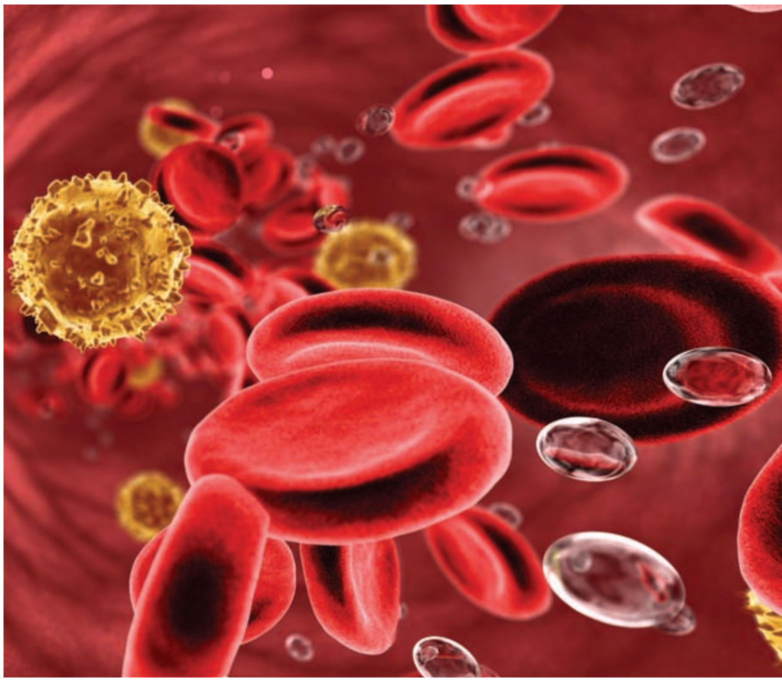
[3] http://www.nytimes.com/2008/06/22/technology/22proto.html?scp=3&sq=gps%20privacy&st=cse accessed Sept 19, 2008

[4] (Jones & Elias, 2006)

possible to use web-scraping technologies to capture up-to-date information on what businesses are doing, rather than relying on administrative records and survey information. Historical records on businesses can also be created by delving into the repository of web pages on the Wayback Machine. This archive takes snapshots of the web every two months and stores them as they were, providing a rich archive of hundreds of billions of web pages. Individual as well as business behaviour can be studied using this archive. Other ways of collecting information on human behaviour from the web include capturing clickstreams from usage statistics. The MESUR project,[5] for example, has created a semantic model of the ways in which scholars communicate based on creating a set of relational and semantic web databases from over one billion usage events and over ten billion semantic statements. The combination of usage, citation and bibliographic data can be used to develop metrics of scholarly impact that go well beyond the standard bibliometric approaches used by academics.

In sum, new approaches to capturing information is transforming the ability of social scientists to do research and to provide information to policy makers. Imagine a similar exercise being done in the study of financial markets, for example. The social science community could potentially transform its empirical foundations if it adopted such a collaborative framework. It could use remote access to a common dataset to move away from the current practice of individual, or artisan, science, towards the more generally accepted community-based approach adopted by the physical and biological sciences. Such an approach would provide the community with a chance to combine knowledge about data (through metadata documentation), augment the data infrastructure (through adding data), deepen knowledge (through wikis, blogs and discussion groups) and build a community of practice (through information sharing). Adopting the type of organisational infrastructure made possible by remote access could potentially be as far-reaching as the changes that have taken place in the astronomical sciences, and cited in the opening section. It could lead to the "democratisation of science" opening up the potential for junior and senior researchers from large and small institutions to participate in a research field.

However, it is worth noting that the establishment of a virtual community to advance the development of a data infrastructure is a social science challenge in its



own right: indeed, the study of virtual organisations is attracting attention as a way of advancing scientific knowledge and developing scientific communities. The report [6] from a 2008 workshop sponsored by the US National Science Foundation shares systematic knowledge about the components, characteristics, practices and transformative impact of effective virtual research environments (VREs); identifies topics for future research that will inform the ongoing design, development, and analysis of VREs for science and engineering research and education; and defines a new cross-disciplinary VRE research community to conduct research across a range of important topics.

A Foundation for More Effective Pursuit of Key Challenges

In Section 2 we asserted that one of the goals of the e-Science Programme was, or should be, to dramatically enhance the ability of the UK to contribute to meeting key or grand challenges facing our world. We listed a set of Grand Challenges articulated by the Research Councils as a whole. The figure below provides a specific mapping between existing e-Science projects and their relevance to capacity building to meet key challenges articulated by

[5] *MESUR: Metrics from Scholarly Usage of Resources http://www.mesur.org/MESUR.html*

[6] *http://www.educause.edu/Resources/BeyondBeingThereABlueprintforA/163051*

the ESRC community (see **http://www.esrc.ac.uk/ESRCInfoCentre/about/delivery_plan/priorities_and_funding/index.aspx**). The figure was presented by Professor Dave DeRoure based on work by Professor Rob Procter and illustrates an analysis of the relevance of some of the current portfolio of e-Social Science projects to these Key Challenges. The Panel has not evaluated this analysis in full but found it accurate for the cases at which we looked. It likely understates the impact in that many other relevant projects are not included. Furthermore, the e-Science Programme is producing generic infrastructure that will facilitate bringing together in distance-independent ways the human expertise, the tools and the data necessary to tackle these challenges.

ESRC Key Challenges	DAMES	DReSS II	LifeGuide	PolicyGrid II	Obesity e-lab	GENeSIS	OeSS II	eStat	CQeSS	MiMeG
Succeeding in the Global Economy	++			+			++	++	+	
Energy, Environment and Climate Change	+		+	++		+	+	+	+	
Education for Life	+			+				++	++	
Understanding Individual Behaviour	+	+	++		++	+	++	++	++	++
Population Change	+	+	+	+	+	++	++	++	+	+
International Relations and Security	+						+			
Religion, Ethnicities and Society	+	+				+	+			+

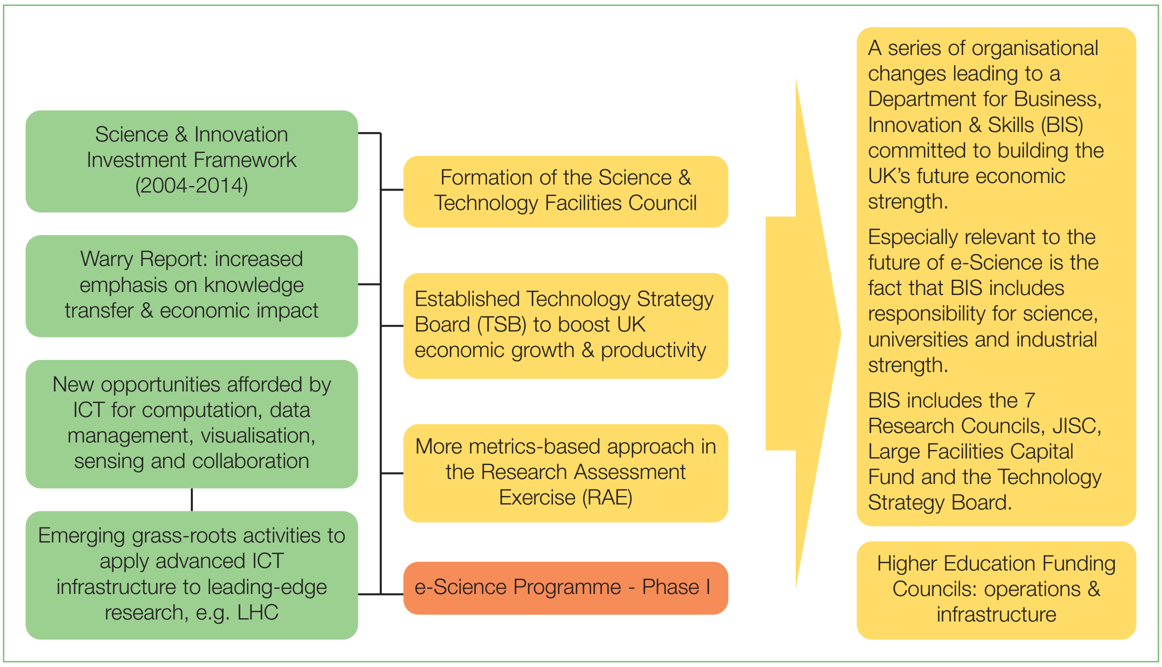
Stretching the Vision and Elevating the Response

The definition of and goals for e-Science established at the inception of the Programme have served the Research Councils and their stakeholders well. But could the platform that has been laid not only scale to broader use, but also move to more transformative impact? For example:

- Might we aspire to create functionally complete virtual research environments (a.k.a. collaboratories) that provide science teams access through the Internet to all the colleagues, all the data, computational tools and observatories required for the project? If so, could they seamlessly work together whether they are located in the same place or different places and interacting at the same or different times?
- How could we do this in agile ways – quickly setting up and taking down a virtual research environment (VRE) in response to new opportunities or threats?
- How could such capabilities increase the intellectual cross-section of ideas coming together to increase the probability of truly breakthrough discoveries?
- How could e-Science methods support more jointly beneficial research interactions with other countries?
- How can we deploy e-infrastructure and environments built upon it in ways that can serve multiple uses: research, education, citizen science, and more effective rapid response to natural or man-made disasters?
- How can services and knowledge from e-infrastructure and e-Science initiatives be applied to learning more generally and especially to more socially based, experiential forms of learning?
- What does the emergence of e-Science and e-humanities & arts say about the future of the research university in the digital age?

With questions such as these, the Panel is suggesting that the UK's pioneering leadership in e-Science is far from completed. It continues to require nurturing and special handling for it has barely begun to achieve its full potential. We are suggesting a process to visit the gains for the past 8 years with the idea of now stretching the vision of possibilities and the ways to get there.

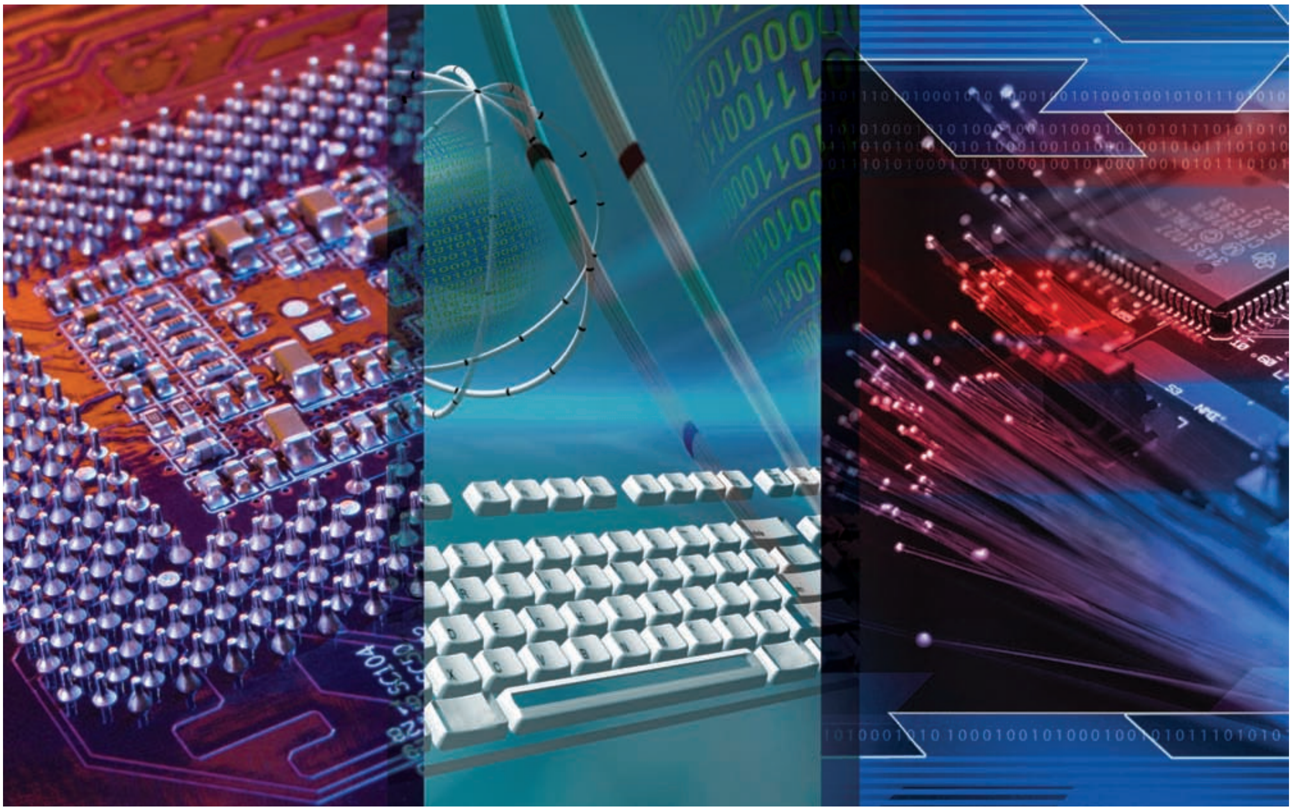
We conclude this section with the figure below which is intended to represent policy and organisational changes described in the Evidence Document that could be relevant to the future of an even bolder e-Science initiative. Although we do not pretend to understand the realities of politics of the formation of the BIS, in principle it contains the remit and a set of organisations in one Department that if appropriately aligned and led, could launch and support an even bolder e-Science, e-research, e-learning initiative in ways that would definitely contribute to building the nation's future economic strength. The UK may at this point be the best positioned nation in the world, including both the US and the EU, to pursue such an effort. Boldly, we would even suggest that the UK has the responsibility to the global science community to continue and expand its leadership in e-Science to produce vision, models, best practices and competition to inform and motivate others.



We note the very important role that JISC has played in piloting and supporting e-infrastructure for the university community and suggest that the JISC role be expanded, better funded and better linked at the highest levels of the Research Councils. The Large Facilities Capital Fund is also highly relevant because e-infrastructure - both hardware, networks and software - are a large, albeit distributed, facility.

We recognise that increased funding for research, or even sustaining the current level, is today a huge challenge, but in the spirit of "no good crisis should go unused" perhaps the UK can take the opportunity for funding reallocation into an even bolder strategic e-Science activity with potentially greater societal benefit across all science fields and education at large.

4. Brief Responses to the Questions in the e-Science Evidence Framework



Introduction

Responses to the following questions are addressed directly or indirectly in the narratives of the previous sections but here we summarise responses to each of the specific questions in the Evidence Framework. The Evidence Document prepared by the review secretariat includes written responses from the e-Science community to these same questions. We will reference these responses in this report, but again stress that the Panel responses are largely based on "primary evidence" gathered through interactions with the community during the review week.

Before writing this report we have carefully studied the comments from the UK e-Science community. The commentary is very insightful and largely consistent with the Panel's analysis. The commentary also reveals that the UK has many researchers in the e-Science community with sophisticated vision, passion and leadership. There are many pioneers working in concert with solid strategic vision. This is a significant asset to be cultivated and used.

Did the UK e-Science Programme build a Platform which enables e-Science tools, infrastructure and practises to become incorporated into mainstream research in the UK?

Did the Programme create a critical mass of capability in developing and exploiting e-Science tools and techniques? Is this capacity being sustained?

There is evidence that significant capability has been created; both in tools and practice. The adoption varies widely across disciplines and projects, and although there is significant impact from the investment to date, there remains much more opportunity for increased adoption and "crossing the chasm" to the mainstream researchers who are not being directly funded by the e-Science Programme. The goal, a long way off, is for e-Science to become the normal way of working. In some specific areas there is critical mass but loss of core funding may put that mass at risk.

Sustainability of research services that have been created is a recurring concern across the entire review process. Sole reliance on business-as-usual grant funding through the individual Research Councils is not likely to take full advantage of the gains to date or the potential for the future.

To what extent have the e-Science technologies developed through the Programme changed the way researchers in other disciplines work? Are there new areas of research that have been enabled by the e-Science Programme? What can be done to increase this happening? Are the present infrastructure and communication channels sufficient?

We have seen numerous examples of e-Science facilitating flows of ideas across disciplinary fields, supporting new ways to extract knowledge from large data collections, augmenting discovery with computational methods, and grounding and informing advances in computer science and engineering. It has stretched the notion of e-Science to be better understood as more than high performance computing. Quoting from the community self-assessment, "High-throughput computing, visualisation, collaborative tools (i.e. social networking for scientists) and data management are now taken more seriously and their relevance to mainstream research is better understood than previously." There are definitely new areas of science being pursued that could simply not be undertaken without the availability, or potential availability, of this more comprehensive form of e-Science.

We saw good examples where workflow and semantic technologies were in use by domain scientists as a way to do their research better. The presence of the grid and its resources clearly has expanded the use of significant computing and data in a number of areas. e-Science has enabled the layering of information (e.g. environmental models, population demographics, future projections, health) in ways fundamental to multidisciplinary practises and knowledge discovery. The shaky mechanisms for supporting tools long-term puts the progress at risk. The physical infrastructure needs significant expansion to support communication and information needs as well as the traditional computational pressures. Long-term sustainability and continuous improvement of the infrastructure is a major challenge that needs to be addressed. One test for a sustainable programme is whether researchers will routinely bet their promotion and tenure, or graduate students will bet their Ph.D., on the future existence of the infrastructure.

To what extent did the e-Science Programme contribute to and benefit from multi-disciplinary research? What barriers to such research did the Programme overcome and what opportunities did it enable?

The e-Science Programme has clearly had major influences in a variety of disciplines, bringing computing experts and researchers in close contact with domain researchers. The barriers are both social (knowing whom to contact and finding mutual benefit) and funding (for the research as well as the facilities and tools). Remaining barriers include career recognition and progress for work that straddles boundaries. There has been significant progress in expanding the scope of "computer science" in the UK as a result of the e-Science Programme. It remains a challenge to establish mutually beneficial collaborations between computer scientists and other domain scientists in ways that both groups are rewarded by their respective fields. A new genre of e-Science professionals are emerging and need to find identity as a field with its own recognition and reward structures.

Although it is not as typical a sentiment as we would like, here is an encouraging quote from the self-assessment concerning interdisciplinary work between computer scientists and application domain scientists:

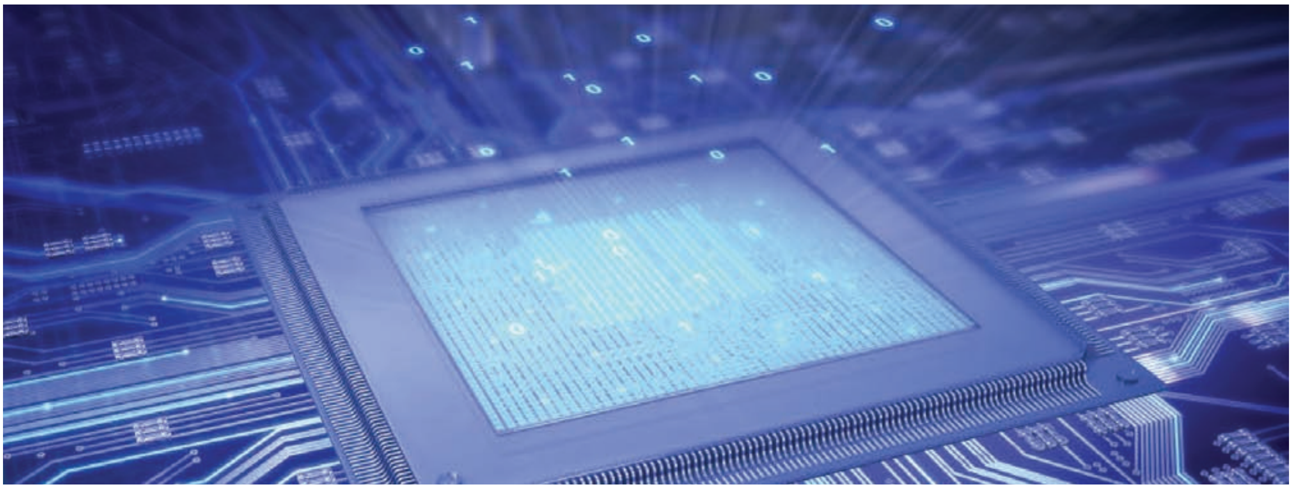
> *The UK Programme built a community and provided the effort needed to facilitate such multidisciplinary engagement. I have been in CS research since the mid 1980s, but I'd not seen anything like this happening before the e-Science Programme. Before the Programme I, like many other computer scientists, often struggled to find real application challenges to drive and evaluate my computing research – I now have regular interactions with researchers from a wide variety of disciplines, and many potential collaborations that can drive our research in a mutually beneficial way.*

To illustrate, however, that the above is not always (yet) true we also include the following quote:

> *Although the Programme succeeded in bringing together computer science groups with other research groups, the collaboration between the two groups led sometimes to misunderstandings: scientists often looked to computer scientists to perform the jobs of software engineers and develop robust tools to order; conversely, computer scientists often – quite reasonably - wished to focus on the latest techniques as part of their own research agenda. The "software engineering gap" continues to be a problem in research computing – who will engineer and maintain the necessary tools? This is an expensive task and engineering work can be seen as taking money away from research.*

The following quote also illustrates success for e-Science facilitating knowledge flow across several application domains:

> *The eSI research themes directly addressed the multidisciplinary barriers by providing a forum for crossing boundaries, leading to transfer of workflows from biology to astronomy, and pooling of scientific gateways across computational chemistry, seismology and cell biology. Thirteen themes in the last five years show the value of this approach, but only begin to address the potential opportunities.*



How effective is the education and training of e-Science practises and techniques at ensuring sufficient take up and adoption? What barriers to effective knowledge exchange and information flow remain and how can they be overcome? What are the barriers to the uptake of these tools and techniques?

The creation of new training and education programmes (MSc and DTC, in expanded and new fields) has made a good start. A few tools also have well-defined and self-supporting training programmes. The National e-Science Centre and JISC run a variety of tutorials and seminars that seem to have been well received. There are still not enough such regular course offerings or high-quality course material. We also agree with the community response that e-Science tools must become fully embedded in undergraduate and post-graduate programmes if the technologies are to become part of the mainstream of research practice, and that we have barely scratched the surface of what is needed for e-Science education and training.

The Panel was surprised and disappointed with the very low participation rate of women in the e-Science Programme. Approximately 97% of the researchers in receipt of grants from the Research Councils in e- Science are men. We suggest that systematically improving this situation, beginning with primary education, be set as a national goal. e-Science, especially through virtual research environments, has the potential to broaden participation in authentic, exciting, motivating science by younger students, their teachers and parents.

It is encouraging to see that there are proportionately more female doctoral students than there are female e-Science Principal Investigators, but we need to realise that this is only part of a complex situation. Even though there are many younger female researchers in the pipeline, it is well-known that a larger fraction of them drop out and don't continue their academic careers. The mechanisms behind this are many and beyond the scope of this report.

How does UK e-Science activity compare globally?

How developed is the global e-Science Platform, and what has been UK e-Science Programme contribution?

Initiatives to create and apply advanced information technology to the conduct of scientific research are underway in most developed and some developing countries under names such as e-Science, cyberinfrastructure-enabled science or cyber science. It is broadly recognised that major investments in infrastructure, its transformative application, and appropriate training are absolutely necessary to remain on the leading edge of discovery and to tackle grand challenge problems confronting society. The UK has been at the forefront of this movement. The Panel believes that the UK e-Science Programme is in a global leadership position in scientific data management (NERC Data Grid), workflow environments (Taverna) and Grid architecture deployment (OGSA DAI). They have done well in e-Science knowledge transfer between fields and between academia and industry. The extensive list of international engagements in section 3.5 of the Evidence Document is impressive.

To what extent did the research undertaken through the Programme engage in "best with best" science-driven international interactions?

The international engagements around e-Science have been generally "best with best." The e-Science Institutes in the University of Edinburgh and the University of Manchester were able to organise several cross-disciplinary courses at an international level for the e-Science community in topics such as managing the Grid e-infrastructure, workflow (Taverna, myExperiment) and distributed Databases (OGSA DAI). The AstroGrid project promotes the adoption of a suite of desktop applications to enable astronomers to explore resources from around the world. The UK e-Science Programme with AstroGrid built a Platform which enables e-Science to be able to use large data collections in Astrophysics around the world. The UK e-Science Programme created the basis for UK scientists to participate in and lead several European projects. This includes collaborations in several disciplines such as Life Science, Bioinformatics, Medicine, Astrophysics, Physics, Social Science and Geoscience.

In which areas of e-Science is the UK the international leader? What has contributed to UK strengths and what are the recommendations for continued strength? In which areas is the UK less strong and what are the recommendations for improvement?

The UK is an international leader in a broad range of areas in e-Science. An analysis of the highlights of the presentation, evidence to the Panel and prior knowledge showed that across the seven disciplines (that is, the Research Councils), over 20 exemplar projects and applications were created that were judged to have had an outstanding international impact. These projects range over high energy physics, the arts, the natural environment (including climate change and oceanography), engineering and health sciences. Besides new tools, new processes and practises enabled by the development of international standards (e.g. Open Geospatial Consortium), the Panel noted increasing interoperability of geographical-based data. The new technologies and tools developed in the frame of this Programme also created the basis for starting a number of small international enterprise initiatives to promote commercialisation and non-academic use.

This success is due to a combination of (1) creating a designated funding stream, (2) implicitly sanctioning e-Science as a critical strategic priority, (3) establishing a cross-cutting management structure to align multi-stakeholder incentives and investments, (4) creating targeted programmes for international engagement, (5) pursuing explicit activities to build community such as the All Hands Meetings, and (6) through all of this, releasing the enormous latent energy of a pioneering research community to pursue e-Science. Continuation of this trajectory is the key to continued success with great focus, in the case of international engagement, on building specific relationships with e-Science activities in other regions of the world including the US (NSF, Department of Energy, NIH), the EU, Australia, New Zealand, Canada and Asia.

The UK e-Science Programme as a whole has yet to create the right environment for sustaining important research facilities that have extraordinary levels of utilisation by their respective communities and high status within their communities. There were a significant number of exemplary services that have been created during the e-Science Programme which have revolutionised the delivery of these services and use and their uptake by the national and international community.

Are there sufficient numbers of research leaders of international stature evident in the UK, in comparison to other countries? If not, which areas have potential for growth?

The e-Science leadership pool within the UK is highly competent, broad and deep – most certainly in the overall size of the community. Many would be readily sought by universities, research institutes and funding agencies in the e-Science area in other countries. As stated earlier, leadership quality is one of the many assets the UK e-Science Programme has helped create that should be nurtured and leveraged into the future.

As mentioned earlier, leadership participation by women has a large potential for growth in numbers but the UK does have several outstanding women in e-Science leadership positions at present.

How does the UK compare internationally with its ability to attract, nurture and support e-Science researchers at every stage of their career?

The UK e-Science Programme created a broad body of e-Scientists with strong discipline and application knowledge. The UK e-Science Programme clearly attracted new scientists to this area, looking to combine e-Science with their application domain. The capacity of the Programme to nurture and support e-Science researchers at every stage of their career has been much more limited. Research in which e-Science is the object of research has had a lower level of support and indeed there is debate about whether this is a research discipline in itself. This disconnect between e-Science and formal or traditional computing science represents a significant risk for ongoing careers of researchers involved in the research around e-Science. There is strong evidence that the hiatus in funds since the end of the UK e-Science core Programme has resulted in e-Scientists being appointed outside of the UK, particularly in the USA. However, there is wide success in the nurturing of application or discipline-specific scientists with an e-Science background who have continued their work in the context of their specific discipline.

The wide support of post-graduate students and early career researchers by the UK e-Science Programme is one of the successes of this Programme, and while the ongoing nurturing of these younger scientists has been mixed, those with a strong discipline-based background seemed the most likely to further their careers through publication. However this outcome raises important questions about how cross-discipline research should be valued in the national context and facilitated across Research Councils, particularly through the facilitation of publication (e.g., journals that support cross-discipline research) and an understanding of impacts that extend beyond simple metrics of journal citations.

In the past eight years how has the UK's global reputation for e-Science research changed?

The reputation has definitely grown, particularly as research communities outside the UK come to realise the enormous strategic importance of e-Science. Some reasons for this growth in reputation include:

- The UK e-Science Programme has taken an active part in developing standards and tools for the national infrastructures for computing and played an important role in several international scientific networks and consortia. These are demonstrating in the participation in many European Networks Infrastructure projects (e.g. EGEE, EGEE-II, EGEE-III, OMII-Europe, GEANT and several other projects to build e-Science applications, standards and training).

- The UK e-Science Programme was able to create the basis for active participation in the European Infrastructure Design Studies in the framework of the European Strategy Forum on Research Infrastructure (ESFRI) programme.

- The UK e-Science Programme was able to create the facility for training a large number of researchers in this field. The capability to organise nationally and internationally several courses for teaching the e-Science technologies created the basis for leading the development in interdisciplinary research and commercial applications.

- Wide awareness of UK e-Science in Australia contributed to the design of e-Research activities and informed the implementation of the National Collaboration Research Infrastructure Scheme (NCRIS) which includes e-Research support for each of the infrastructure areas.

- The pioneering activities of the UK in e-Science helped stimulate the launch of similar activities elsewhere including in the US.

What has been the impact (accomplished and potential) of the UK e-Science Programme?

To what extent has the research undertaken through the e-Science Programme benefited the UK economy and our global competitiveness?

The Panel found that a surprisingly large fraction of the presented projects had in one way or another generated recognisable commercial impact. We were presented with illustrative examples covering a wide range of commercial impact mechanisms, such as creation of spin-off companies, transfer of knowledge to existing companies, and active commercial participation in e-Science research and development projects. Benefits to commercial partners could be seen in terms of development of new methods, access to infrastructures providing data storage and processing facilities, and perhaps most importantly enabling competence transfer and a basis for recruitment of skilled human resources.

Section 3.7 of the Evidence Document contains significant evidence of knowledge transfer and collaborative research with industry. We judge the participation of the Department for Trade and Industry, the Technology Strategy Board and in the future major units of the new BIS to be a significant strength of the UK e-Science strategy.

The Panel noted that that bridges between universities and large as well as small companies have been built with the DAME project as a prime example of the involvement of a large company. Presenters from Newcastle made a compelling case for the regional economic impact of the e-Science Programme. The Panel did, however, get the impression that the underpinning mechanisms generating the commercial impact have to a large extent depended on initiatives taken by individual projects. In many cases it appeared that the industrial partnerships had evolved before, or outside of the framework of the e-Science Programme, but had found an additional breeding ground within the e-Science projects.

A robust e-Science Programme, underpinned by an advanced e-infrastructure, is and will continue to be essential for leading edge discovery, innovation and collaboration on a global scale. It will provide a competitive advantage in a global, knowledge-based economy. The UK has a great head start in this direction.

It is also noteworthy that Professor Tony Hey, the Director of the inaugural e-Science Programme, was recruited by Microsoft essentially to nurture an e-Science activity at Microsoft. This highly visible move is evidence of the international visibility and high standing of the Programme, but in addition sends signals to the world that e-Science has commercial potential both in provisioning and application.

To what extent did the research undertaken through the UK e-Science Programme address key technological/societal challenges?

Section 3.3 of the Evidence Document includes details about the e-Science projects funded by the various research councils in concert with the core funding. The general impression, confirmed by the Panel's interaction with about 60 projects at the Review, is that the UK has built a significant portfolio of important projects addressing major technical and societal challenges in the physical, biological, social and information sciences.

The UK research community furthermore realises that the key to addressing the most significant societal challenges resides in the complementary expertise, data, tools and facilities of multiple disciplines and that e-Science is critical to the required synergistic merger. The e-Science Programme has created a foundation that can be built upon to facilitate addressing all of the Grand Challenges of the RCUK mentioned in Section 2 as well as, for example, research areas recently

ESRC Key Challenges	DAMES	DReSS II	LifeGuide	PolicyGrid II	Obesity e-lab	GENeSIS	OeSS II	eStat	CQeSS	MiMeG
Succeeding in the Global Economy	++			+			++	++	+	
Energy, Environment and Climate Change	+		+	++		+	+	+	+	
Education for Life	+			+				++	++	
Understanding Individual Behaviour	+	+	++		++	+	++	++	++	++
Population Change	+	+	+	+	+	++	++	++	+	+
International Relations and Security	+						+			
Religion, Ethnicities and Society	+	+				+	+			+

presented by the ESRC (see **http://www.esrc.ac.uk/ESRCInfoCentre/about/delivery_plan/priorities_and_funding/index.aspx**).

The figure above, based on work by Professor Rob Procter, and mentioned earlier in the report was presented by Professor Dave DeRoure and illustrates an analysis of the relevance of some of the current portfolio of e-Social Science projects to these Key Challenges articulated in the economics and social science arena (these overlap, but are not the same as the RCUK Grand Challenges mentioned earlier). The figure likely understates the impact of the e-Science Programme to date in that many other relevant projects are not included and there are other Grand or Key Challenges not included in the table. To be sure though, the e-Science Programme is producing generic infrastructure that will facilitate bringing together, in distance-independent ways, the human expertise, the tools and the data necessary to tackle challenges on a global scale.

We note that the e-Science Programme had a number of significant impacts on health care research and practice. It is recognised that the main impact on the health care system is not in the e-Science projects as such, but rather in the discoveries that they will enable through the use of the tools that are developed. One project, CancerGrid, which supported clinical trials, had already accelerated clinical trial research and saved more than the investment cost.

The ultimate impact is hard to predict due to the long timescale at which these results migrate into clinical routine. It is, however, clear to the Panel that many results having potentially large impact have been, and will be, generated based on the developed e-Science platform. Many high-quality examples of this are found among the projects in which genomic and proteomic data have been collected and stored and advanced analysis tools have been developed. Other significant impacts on health services can be found in the efforts made to use e-Science approaches for patient records, including patient data collection, curation and mining. This creates opportunities both for health researchers to conduct more elaborate studies of larger populations and to improve and personalise health care.

Other areas of societal impact are in the environment and climate sectors. Advanced environmental planning is now using sensor data, simulations and variety of data sources increasingly made interoperable through data standards. An example of a project in this area is the ClimatePrediction.com project which makes use of grid technologies to allow the general public to contribute to climate change simulation by volunteering their personal computers' free cycles. This project is also an example of how an e-Science project can have an impact by encouraging the public to form opinions and contribute to the climate change debate. Several projects showing similar effects engaging the general public by maintaining interesting homepages accessible to the general public were noted in many different application areas. There are also examples of projects that have the goals of generating content for public spaces such as museums and science centres.

One of the impacts, touched upon earlier, is the building of social networks and communities of work practice. With very few exceptions, all the project managers noted the importance of the contacts established with scientists in other fields. In most cases these contacts have prevailed after the completion of the projects. The core e-Science funding mechanisms have been catalytic for building collaborations and enabling new research directions. The All Hands Meetings were mentioned as one of the key events for the establishment of these networks.

One of the most valuable forms of impact and knowledge transfer is the training of skilled, e-Science trained professionals. The primary resource is the PhD students working on the projects who, after completion of their training, find their employment outside of academia.

The vision for e-Science is to enable the seamless integration of data, computation and research practise (or workflow). The UK e-science Programme has had a very significant impact in the development of standards in all RCUK Council areas. These standards developed in the framework of international and national communities with established practises that facilitate the exchange of data within and across disciplines. It is these practises and standards that facilitate interoperation and data sharing that builds real capacity for new science now and into the future.

What evidence is there to show that the UK e-Science Programme supported the development of a creative and adventurous research base and portfolio?

The strongest evidence for this is the large set of statistical evidence and thoughtful community narrative in the Evidence Document, reinforced by our findings from the intense week of mostly high-quality, stimulating and enthusiastic interactions with the community. The Panel worked very hard – but enthusiastically – precisely because of the creative and adventurous things they were hearing and the palpable enthusiasm of the community.

Examples of diverse, adventurous and creative e-Science activities, many funded at a low level, became very clear during the week. Examples include projects in archaeology, dance and human movement, text recognition, bus timetabling, clinical trials in health, remote control of instruments. These were in addition to systems tailored for scientists to collaborate in more traditional e-Science application areas such as high-energy physics and climate sciences. It was noted by the Panel that other countries do not have the same diversity and that frequently their e-Science application domains were more narrowly defined.

What are the future opportunities for UK e-Science?

Is the research community appropriately structured to respond to current and emerging technological and societal challenges? If not, what improvements could be implemented?

What are the current strengths, weaknesses and opportunities in UK e-Science?

Where should the Research Councils focus support for e-Science activity in the future?

The pioneers of e-Science have built a community, e-Science centres and training opportunities and have contributed to the creation of e-infrastructure and its adoption. They have created a microcosm of an e-Science environment that can be enhanced and scaled into the future. The best of these activities need to be identified and supported in a competitive way to move forward in innovation and to attract and support more practitioners – to cross the chasm we described in Section 3. The readiness of others to join still varies enormously and variable-rate take-up needs to be appreciated as the norm. The pioneer community is connected in effective networks in two complementary dimensions: by institutions and regions; and by disciplines and projects. These networks have international reach. The community seems to have a healthy balance between cooperation and competition.

Further success will not be achieved purely through technological determinism nor through the uncoordinated activities within the various disciplinary communities. Moving forward requires coordination, clever design, effective leadership, and long-term commitment to a system of linked and balanced interaction between the various communities and sponsors through:

- The creation, enhancement and sustaining of application-driven, shared e-infrastructure and comprehensive services developed collaboratively between technologists and users using iterative design methods.

- The alignment of a variety of stakeholders (Research Councils, JISC, Large Facilities Capital Fund, Technology Strategy Board, Higher Education Funding Councils, universities, private philanthropy, industry and international research funding agencies) to jointly co-funded, in ways consistent with their various missions, a sustained and continuously improving e-infrastructure as a platform for e-Science;

- The funding of the research communities, particularly those relevant to grand/key challenges, to use and benefit from the e-Science environment in a high-quality and transformative way.

- The funding of technical and social science research communities for which e-Science is an object of research.

- Programmes for the systematic transfers of the fruits of both types of research listed above into innovative systems and processes that benefit all facets of society, including the environment and the economy.

- Incentive structures to reward joint work within and between e-Science application fields and the fields of computer, information and social science relevant to creating the tools and resources.

- Educational and training mechanisms to enhance human capacity to both create and use e-Science environments.

- Mechanisms for the UK research community to routinely participate in, and often lead, global-scale science projects (this requires attention to an international, interoperable e-Science environment).

The comprehensive services mentioned above need to include integrated high-performance computing, data stewardship and access, visualisation and human interaction capabilities, scholarly communication environments, virtual research environments and online instrumentation. Processes need to be established to guide and sustain the creation of a **shared** e-infrastructure on which project-specific software can be easily tailored. A focus on sharing is not driven only by cost efficiencies (which can be very significant) but also by the need to support interoperability between projects and disciplines. Global scale interoperability is also very important. An uncoordinated approach to e-infrastructure could lead to balkanisation of e-Science just as the need for interdisciplinary research becomes more imperative.

Strength - The UK has created a "jewel" – a pioneering, vital activity of enormous strategic importance to the pursuit of scientific knowledge and the support of allied learning.

Weakness - The danger is that the UK's jewel of e-Science, and its current competitive advantage relative to the rest of the world, will be lost through equivocal support at the national level. There are dangers in dissolving a well-defined and vigorously led e-Science activity that cuts across all the Research Councils (a matrix organisation) and that has resources to leverage against the core resources of the various Councils. There is need for greater involvement with the HPC community and more systematic interaction with industry. There is need to establish models for long-term funding and support of critical components of e-infrastructure: networking, HPC, data and information repositories, open software development and hardening, and training programmes to build human capacity to use and support e-infrastructure. There is need for a strategy to span the chasm between early adopters and the mainstream research community.

The generally downward trend of e-Science project and contract project count and funding in total, and by individual Research Councils, shown in Figures 8 through 14 of the Evidence Document is very worrisome. The upward trend in funding by the STFC, shown in Figure 15, is encouraging but far from sufficient and it does not balance infrastructure investment with research investment.

We conclude with a repetition of the observation we made at the end of Section 3:

Although we do not pretend to understand the realities of politics of the formation of the BIS, in principle it contains the remit and a set of organisations in one Department that if appropriately aligned and led, could launch and support an even

bolder e-Science, e-research, e-learning initiative in ways that would definitely contribute to building the nations future economic strength. The UK may at this point be the best positioned nation in the world to do so including both the US and the EU. Boldly, we would even suggest that the UK has the responsibility to the global science community to continue and expand its leadership in e-Science to produce vision, models, best practices, and competition to inform and motivate others.

We note the very important role that JISC has played in piloting and supporting e-infrastructure for the university community and suggest that the JISC role be expanded, better funded and better linked at the highest levels of the Research Councils. The Large Facilities Capital Fund is also highly relevant because e-infrastructure - both hardware, networks and software - are a large, albeit, distributed facility.

We recognise that increased funding for research, or even sustaining the current level is today a huge challenge, but in the spirit of "no good crisis should go unused" perhaps the UK can take the opportunity for funding reallocation into an even bolder strategic e-Science activity with potentially greater societal benefit across all science fields and education at large.

How did the Programme Strategy (having a Core and individual Research Council Programmes, developing tools and applications in parallel) affect the outputs from UK e-Science?

What progress would have occurred if a specific Programme had not been in place?

What was the impact of the Programme on the provision of skills and trained people in the UK?

What was the 'added value' of the Programme strategy? Is this a good model for cross-council Programmes? Could the model be improved and if so it what way?

There is broad agreement within the Panel and, for our conversations and reading also within the UK e-Science community, that the organisational and management strategy has worked well. There is also agreement that the e-Science Programme should not be viewed as a one-shot, five year project, but rather as a long-term strategy that needs to be continuously refined and carried forward indefinitely.

It is difficult to say what would **not** have happened without this Strategy, but we speculate that the UK would not have received such a high return-on-investment in both tangible, intangible and potential assets were the funds allocated in the normal way through Research Councils. The total e-Science investments over the first 5 years were less than the cost of one petascale computer such as the NSF-funded Bluewaters in the US (the UK does need to plan for access to, not necessarily ownership of, such machines). We doubt that much of the core technology development would have been possible to fund through the individual Research Councils, and if it were found, may have led to duplicative, perhaps even stove-piped, services.

Strengths of the strategy to date include:

- visionary and effective leadership that was given a good balance between remit, authority and resources;
- a focus on data and collaboration aspects of e-Science, and not only on high-performance computing for modelling and simulation;

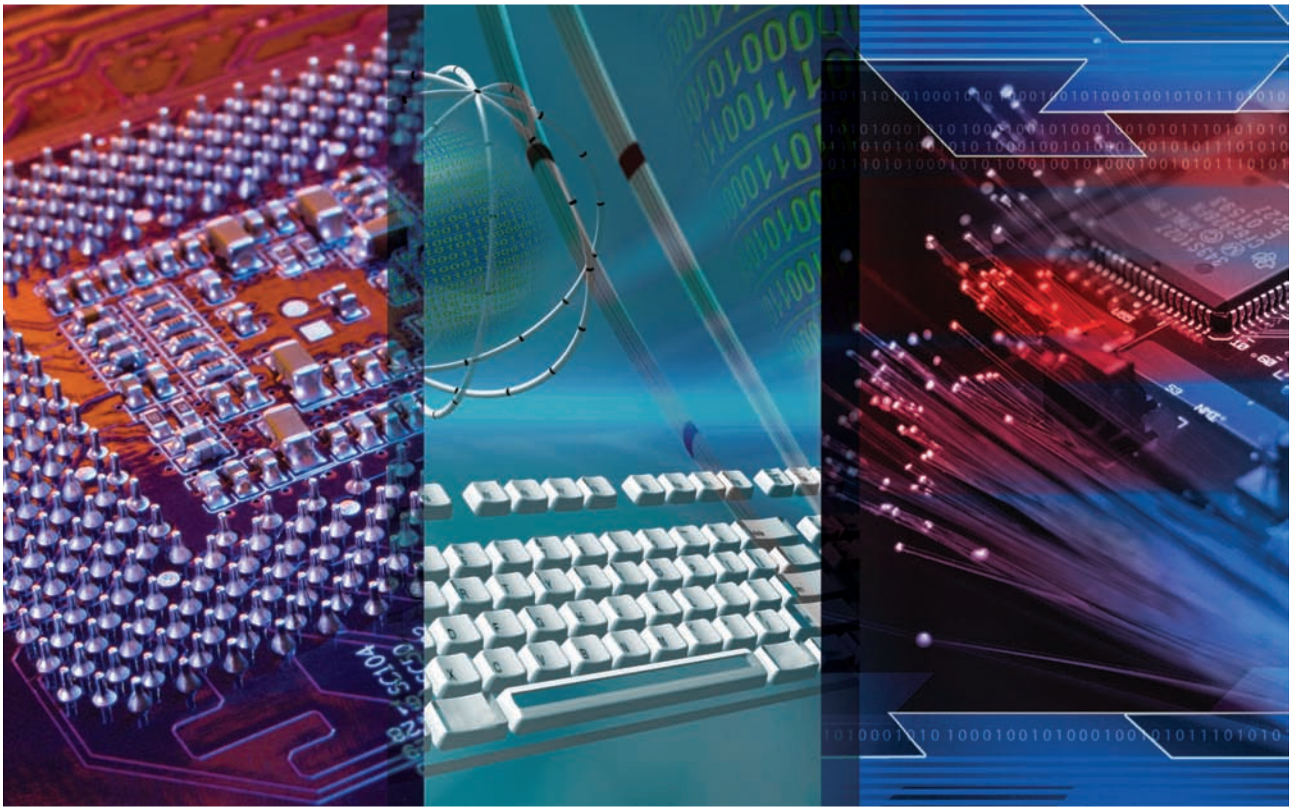
- the participation of the DTI to provide both industrial contacts and complementary, flexible funding;
- a well-judged balance between investments in pioneering research and delivery of infrastructure and generic services (we read one comment that the Programme was too infrastructure heavy but the Panel does not agree);
- a crosscutting funding and management platform to help build interdisciplinary communities;
- a conscious effort to attract key research leaders into the Programme early; and
- a fast ramp-up (but a weakness is a fast ramp-down).

A more permanent crosscutting Programme with real authority and resources continues to be important. We are suggesting a hierarchy of small but energetic, value-adding coordinating and leveraging organisations for e-Science along the following lines:

- an office to coordinate relevant offices within BIS as discussed earlier;
- an office to coordinate and leverage across the Research Councils;
- a network of regional centres coordinating and leveraging across their portfolio of projects.

International and industrial engagement and applied educational needs and opportunities should be addressed at all these levels.

5. Major Conclusions and Major Recommendations



Major Conclusions

The e-science movement has emerged from a combination of push and pull. It is propelled by the push of the exponential growth of ICT capabilities, coupled with the pull of the demand for transformative tools and methods now needed to support the complexity, diversity and integrative needs of modern and future scientific research. The fundamental goal of the global e-Science movement is to determine how to use ICT as a foundation (as e-infrastructure) for transformative enhancement of the *doing* of research, in ways that create more transformative benefit from the *results* of research. e-Science is about transforming knowledge discovery in science; it is about innovation to support innovation.

The path linking research, knowledge production and innovation is complex and nonlinear but investments in e-Science are critical for expanding the knowledge creation that lies at the heart of innovation. Innovation is fundamental to advancing economic and social well-being. Furthermore, investments in e-Social Science offer the potential to better understand those pathways and to fully turn the UK's investments in basic research into advances in innovation and economic prosperity, including high wage jobs. The e-Science Programme can also be a pilot to show the way towards ICT-enabled environments for more effective and inclusive, anytime and anywhere, life-long learning.

The technologies and practices of e-Science, together with the e-infrastructure on which it rests, must be both a topic as well as an enabler of research and development; and this duality needs to be made synergistic. e-Science as a topic of research includes both technological and social (behavioural, economic, legal, ethical) dimensions. As an enabler of research it requires establishing and nurturing mutually-beneficial relationships between those skilled in design and evaluation of e-science environments, those pushing the edge using these environments and those providing operational services and training. It is intrinsically an interdisciplinary, multi-role team effort.

The Panel has concluded that the UK e-Science Programme is in a world-leading position along the path described in the title of this report: *Building a UK Foundation for the Transformative Enhancement of Research and Innovation*. The investments to the present are already empowering significant contributions to well-being in the UK and the world beyond. The UK must now decide whether to create the necessary combination of financial, organisational and policy commitments to capitalise on their prior investments, and to move to the next phase of building capability, growing adoption and achieving competitive advantage.

The successful creation and adoption of e-Science is an organic, emergent process that requires ongoing, coordinated investment from multiple funders together with coordinated action from multiple research and infrastructure communities. It requires nurturing robust infrastructure and a continuous cycle that couples research, application development and training processes. It is the balance between these processes that drives success in e-Science. None of this is easy but the rewards for success are enormous.

Major Recommendations for Action

The previous sections of this report contain interwoven findings and recommendations, as well as direct responses to the questions in the Evidence Framework in Section 4. We have given both a retrospective assessment in Section 2 as well as our vision of opportunities and challenges for the future in Section 3. The UK e-Science community has also given the RCUK the benefit of their assessments and advice for the future of the e-Science Programme in Section 4 of the Evidence Document. All of this, we hope, will be studied carefully by the e-Science community and the relevant funders and be helpful for plotting the future of e-Science in the UK.

We conclude this report with a list of a dozen major recommendations for action at a more general level than in earlier sections. The emphasis here is on *what to do*, rather than being very prescriptive about *how to do* it. These are not in a priority order, but most are quite interdependent.

1. Structure and leadership
Establish organisation and management structures that continue to treat e-Science as a designated strategic initiative spanning all Research Councils and having ongoing designated funding. Provide high-quality dedicated leadership for a strategic e-Science Programme, and provide the leader with adequate authority and resources to catalyse real synergy within and between funder, researcher and service providers. The leader needs resources to co-fund with specific projects funded by the individual Councils. This recommendation includes exploiting more systematic, coordinated investments in the infrastructure, development, and adoption of e-Science between several components of BIS, including the Research Councils, JISC, the Technology Strategy Board and the funders of higher education and facilities. e-Science Programme leadership should also seek coordination with e-Science-related funding from outside of government including the Wellcome Trust, the EU and others.

2. Industry-academic collaboration
Establish more systematic and better supported mechanisms, including targeted funding, to nurture collaboration and bi-directional knowledge transfer between academia and industry in the creation, provisioning and application of e-Science. There are multiple goals here: (1) to identify better ways of connecting academia and industry through e-Science to accelerate the transformation of research outputs to beneficial innovation; (2) to enlist relevant industry in the creation and perhaps provisioning of the e-infrastructure platform for e-Science; and (3) to help industry adopt and tailor best practices and services from e-Science to enhance their own productivity.

3. RCUK e-Science Centre network
Sustain and strengthen the RCUK network of e-Science Centres. Challenge and support these Centres to both serve their regional constituencies and be members of a network for the common good with others in the UK and international partners. Establish the remit of various centres in a way that stresses complementary expertise and sharing. Emphasise and assess expectations that the Centres be proactive in engaging with each other, with UK industry and with other countries. Establish polices for sustaining this network but in a way that periodically (e.g. every five years) requires re-competition to enable new entries into the field and to retire less effective activities.

4. Sustaining advanced e-infrastructure
Sustain the operational e-infrastructure for e-Science created to the present. Informed by an ongoing review of future needs, evolve it to: (1) higher capacity; (2) more complete function; and (3) leading-edge, distributed system architecture. E-infrastructure, as we are using the term, includes the hardware, software, organisations and people to provide generic but tailorable services such as networking/communication, computation, visualisation, data repositories, digital libraries, online observatories, sensor networks and instruments, and distributed (virtual) collaboration. Middleware provides the glue to integrate these services in secure ways. In particular the UK needs to invest in e-infrastructure in ways that (1) anticipates the continuing exponential growth in scientific data and the increasing ability to extract knowledge from it; (2) provides the UK research community access to petascale-level computing and anticipates the future needs for access to exascale; (3) continues to build on and strengthen the grid model of distributed computing, but also explores the adoption of emerging models of cloud computing for research. Scaling scientific computer codes to the peta- or exa-level is generally a major challenge that would require R&D support.

5. Supporting complementary roles
Recognise in programme calls and funding policies that there are people in several complementary roles that need to be funded in a balanced way. There are (1) researchers seeking to innovate in the application of e-Science methods; (2) researchers in computer science and some aspects of social science contributing to designing better e-Science methods and services; (3) professional software engineers or informatics specialists who build reliable production-grade systems; and (4) professionals who administer and operate the supporting e-infrastructure. There are people effectively spanning several of these roles. There may be the need to better define and reward new professional identities and job types within academia that span role 2 and 3 above.

6. Sharing for cost and science effectiveness
Continue funding policies that strongly encourage or require the creation and adoption of shared e-infrastructure. This is important not only from a cost-effective, efficient-energy use and environmental-impact perspective, but also for facilitating intellectual interoperability between disciplines, institutions, facilities and data resources essential for many grand-challenge research endeavours. Since scientific research is intrinsically global, place great emphasis on creating UK e-infrastructure that harmonises with the e-infrastructure in other countries. Doing so will

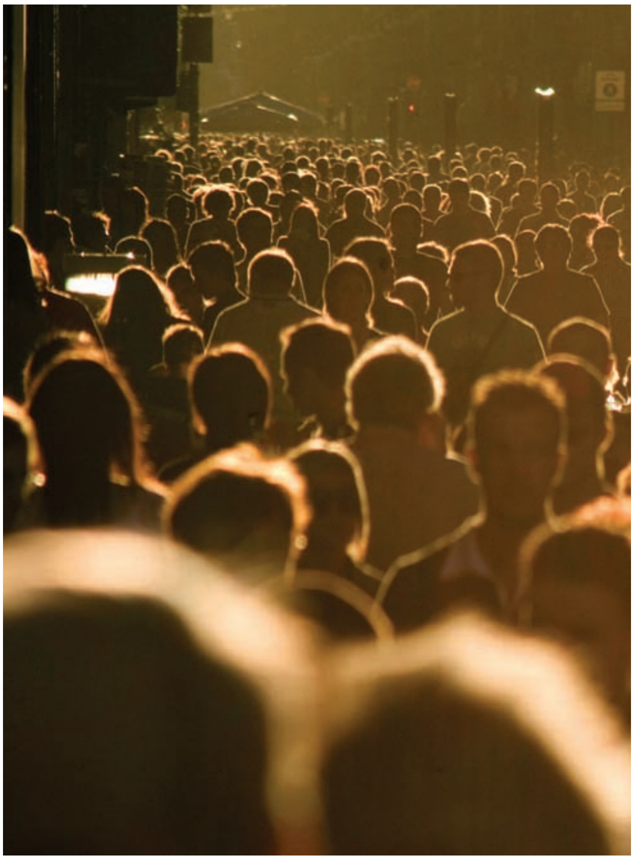
enhance the sharing of data and unique, expensive instruments and will reduce constraints on collaboration at a global scale.

7. Role for arts and humanities
Encourage and support even more participation of the arts and humanities research communities in the e-Science Programme (we saw some excellent beginnings in our review). Arts and humanities are poised to achieve large benefit from e-science

methods and infrastructure as the human record becomes increasingly digitised and multimedia. For example, a field called "corpus computing" is emerging due to the ability now to compute across enormous collections such as those being created by industrial partnerships with academic libraries. Copyright management for such work will remain an issue but for public domain and open-license materials, the field is now wide open.

8. Role for social sciences

Building on a strong start, encourage and support even greater leadership by the social science research community in the adoption of e-Science methods particularly, for example, given capabilities to explore enormous data sets, analyse social networks, and explore very complex systems through simulation and modelling. The social science community should also be encouraged to contribute more to deeper understanding of more principled ways to design effective virtual research environments, collaboratories and four-quadrant environments (see last section below). Many science communities are creating such distributed knowledge communities, but many are sub-optimal or outright failures, usually for social and behavioural rather than technical reasons. If research is conducted within such technology-mediated environments, we can potentially capture and later mine not only the artefacts of knowledge work but also the processes.



9. Crossing the chasm; refreshing innovation

Develop a dual strategy that both (1) accelerates the adoption of e-Science methods in the "mainstream market" of researchers as discussed in Section 3 ("crossing the chasm"); and (2) refreshes the investments in the "early market" to produce the next wave of innovation in e-Science services and application. Goal (1) involves training, making services available, and tailoring current services to more specific needs of disciplinary and interdisciplinary communities. This process needs an integrated formative assessment activity that includes continuous monitoring and that helps ground and inform the activities of "early market" communities. The assessment should inform a spiral, iterative design process. Also include special opportunities for upper management of universities, government and industry to learn more about the fundamentals and strategic importance of e-Science and e-infrastructure.

10. Data stewardship at enormous scale

Continue the strong focus on creating practices and services for appraisal, curation, federation and long-term access to scientific data. Complement or broaden the activities of the Digital Curation Centre with the creation of coordinated and sustained production services for curation and stewardship of scientific data. Consider a highly centralised, large data centre model for storing the information and preserving the bits, together with a distributed model for curation by disciplinary specialists. The academic digital libraries centres might be encouraged to assume some major responsibility for scientific data. Seek and promote international cooperation. In all of this, plan for continued exponential growth in scientific data.

11. Openness as a general policy

At every opportunity establish and support policies for openness: open-source code, open data and open courseware. To the extent possible, these should be freely available with terms of use that encourage reuse. Work with international standards activities especially for interoperable data.

12. Towards functionally complete, four-quadrant, research environments

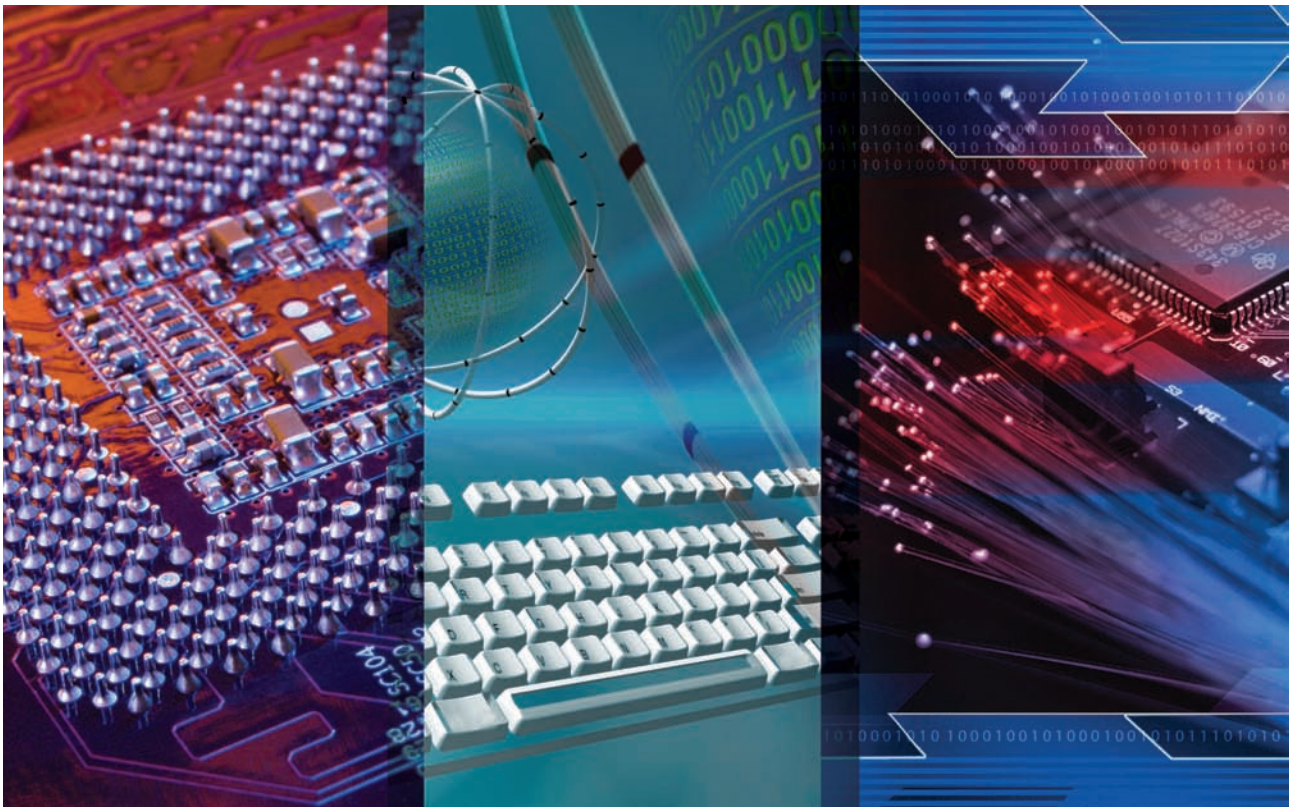
Place greater emphasis on the overarching goal of establishing capacity for collaborative, international, interdisciplinary team science to occur routinely in "functionally complete, four-quadrant environments" built upon e-infrastructure. "A four-quadrant environment" refers to a blended virtual-physical environment in which the activities of a group can flow easily between all four quadrants in a 2-by-2 matrix with same versus different for the two dimensions of both time and place. It subsumes the concept of virtual research environment. "Functionally complete" means that the environment supports access to all the people, the information and data services, the observatories and facilities, the computational services, and the collaboration and communication services necessary for a scientific team (or more generally, a community of practice) to carry out its work. Such environments could become both necessary and sufficient for participation in a research endeavour. They could accelerate and broaden participation in scientific discovery and learning. They offer the potential to support both explicit and tacit knowledge creation, and thus to support a blend of learning *about* science, learning *to do* science, and learning *to be* a scientist.

Geographic Place	Same (synchronous)	Different (asynchronous)
Same	**P:** Physical mtgs **I:** Print-on-paper books, journals **F:** Wet labs, studios, shops	**P:** Shared physical artifact **I:** Library reserve **F:** Time-shared physical labs, ...
Different	**P:** AV conference **I:** Web search **F:** Online instruments	**P:** Email **I:** Knowbots **F:** Autonomous observatories

P: people, I: information, F: facilities, instruments

Four Quadrant Organisations (virtual + physical organisations) offer additional modes of interaction between People, Information and Facilities

Annexes



PROFESSOR DANIEL ATKINS

Professor Daniel E. Atkins is a Professor in the School of Information and in the Department of Electrical and Computer Engineering at the University of Michigan (UM), Ann Arbor. Beginning September 2008 he became the W. K. Kellogg Professor of Community Informatics. Effective 1st September 2008 he also accepted a part-time position as the Associate Vice-President for Research Cyberinfrastructure at the University of Michigan. From June 2006 to June 2008 he was on leave from the university to serve at the National Science Foundation as the inaugural Director of the Office of Cyberinfrastructure.

He began his research career in the area of computer architecture and did pioneering work in parallel computer architecture and high-speed computer arithmetic that is widely used in modern processor chips. He has served as Dean of the College of Engineering and as the founding Dean of the School of Information at the University of Michigan, Ann Arbor. The School has been a catalyst for creating an international Information School (I-school) movement. He was founding Director of the Alliance for Community Technology (ACT), an international partnership with philanthropy for research and development in the use of information and communication technology (ICT) to further the mission of educational and other non-profit organisations.

Professor Atkins does research and teaching in the area of distributed knowledge communities and open learning resources. He has directed several large experimental digital library projects as well as projects to explore the socio-technical design and application of "collaboratories" for scientific research. He served as Chair of the National Science Foundation Advisory Panel on Cyberinfrastructure. The Panel issued a landmark report in February 2003 recommending a major Advanced Cyberinfrastructure Programme intended to revolutionise science and engineering research and education. The report has catalysed new priorities and the new Office of Cyberinfrastructure at the NSF.

Professor Atkins also serves regularly on Panels of the National Academies exploring issues such as scholarship in the digital age, the future of scholarly communication and the impact of information technology on the future of higher education. He is co-author of Higher Education in the Digital Age: Technology Issues and Strategies for American Colleges and Universities. He has served as an international consultant to industry, foundations, educational institutions and government. His recent report, A Review of the Open Educational Resources (OER) Movement; Achievements, Challenges, and Opportunities, with J. S. Brown and A. L. Hammond is now helping to shape international investment in the next phase of the open courseware movement.

Professor Atkins was the 2008 winner of the Paul Evan Peters Award from the Coalition of Networked Information, Association of Research Libraries and EDUCAUSE. The award recognises notable, lasting achievements in the creation and innovative use of information resources and services that advance scholarship and intellectual productivity through communication networks. Previous award recipients include Paul Ginsparg (2006), Brewster Kahle (2004), Vinton Cerf (2002) and Tim Berners-Lee (2000).

In May 2009 he was recognised with a University of Illinois College of Engineering Distinguished Alumni Award for his influence on high-performance computer architecture, pioneering work in the development of schools of information and leadership in improving the U.S. cyberinfrastructure.

PROFESSOR NATHAN BINDOFF

Professor Nathan Bindoff is Professor of Physical Oceanography at the University of Tasmania, and CSIRO Marine and Atmospheric Research (CMAR) Laboratories, Director of the Tasmanian Partnership for Advanced Computing (TPAC) and Project Leader of the Antarctic Climate and Ecosystems Cooperative Research Centre's (ACE CRC) Modelling Programme. He is a physical oceanographer, specialising in ocean climate and the earth's climate system.

Professor Bindoff has been the coordinating lead author for the ocean chapter in the Inter-Governmental Panel on Climate Change (IPCC) Fourth Assessment Report. Together with his colleagues, they have documented some of the first evidence for changes in the climate change signals in the Indian, North Pacific, South Pacific and Southern Oceans and shown some of the first evidence of changes in the

Earth's hydrological cycle. He has established the programmes and experiments that determined the total production of Adelie Land Bottom Water formation and its contribution to Antarctic Bottom Water Formation, contributed to the development of some of the largest and highest resolution model simulations of the oceans and has been deeply involved in oceanographic data and data management as the chairman of the Data Products Committee for the World Ocean Circulation Experiment and the International Polar Year. TPAC operates a federated digital library in support of the Earth Systems Science community in Australia and develops software to enhance connectivity between climate and geospatial communities. In his spare time he has led 9 Oceanographic voyages on the Aurora Australis in the Southern Ocean.

PROFESSOR CHRISTINE BORGMAN

Professor Christine L. Borgman is Professor and Presidential Chair in Information Studies at UCLA. She is the author of more than 180 publications in the fields of information studies, computer science and communication. Both of her sole-authored monographs, Scholarship in the Digital Age: Information, Infrastructure, and the Internet (MIT Press, 2007) and From Gutenberg to the Global Information Infrastructure: Access to Information in a Networked World (MIT Press, 2000), have won the Best Information Science Book of the Year award from the American Society for Information Science and Technology. She is a lead investigator for the Centre for Embedded Networked Systems (CENS), a National Science Foundation Science and Technology Centre, where she conducts data practices research. She chaired the Task Force on Cyberlearning for the NSF, whose report, Fostering Learning in the Networked World, was released in July, 2008. Professor Borgman is a Fellow of the American Association for the Advancement of Science (AAAS).

She is currently a member of the U.S. National Academies' Board on Research Data and Information and the U.S. National CODATA (Committee on Data for Science and Technology), the Strategic Advisory Board to Thomson-Reuters Scholarly Research and the Advisory Board to the Electronic Privacy Information Centre. She is Member-at-Large for Section T (Information, Computing and Communication) of the AAAS. At UCLA, she chairs the Information Technology Planning Board. Professor Borgman is a member of the editorial boards of the Journal of the American Society for Information Science & Technology, The Information Society, Journal of Digital Information, International Journal of Digital Curation, Information Research and the Journal of Library & Information Science Research.

PROFESSOR MARK ELLISMAN

Professor Mark Ellisman is Professor of Neurosciences and Bioengineering, Director, Centre for Research in Biological Systems, University of California, San Diego, La Jolla, California.

Professor Ellisman is an expert in the development and application of network and information technologies to advance the biological sciences. His research furthers investigations in the basic molecular and cellular mechanisms of the nervous system and enables the development of advanced technologies in microscopy and computational biology. He is a pioneer in the development of three-dimensional, light and electron microscopy and the application of advanced imaging technologies and computational resources to achieve greater understanding of cellular structure and function, particularly applied to the nervous system. He has co-authored several hundred publications in the fields of neuroscience, labelling technologies for molecular imaging, development of advanced research instrumentation, image processing, visualisation and knowledge management.

Professor Ellisman's telemicroscopy research initiative predated the advent of the web browser and was instrumental in the evolution of cyberinfrastructure, i.e. the use of advanced networks to connect computational, data storage, visualisation and software tools with rare research assets to address global research priorities.

He recently expanded his efforts to build community resources using emerging capabilities in information technology to two new biology initiatives: a project in neuroscience to create an electronic "Whole Brain Catalog;" and a community-serving cyberinfrastructure for microbial metagenomics known as CAMERA (Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis).

In 2001, Professor Ellisman launched the development of the National Institutes of Health (NIH) Biomedical Informatics Research Network (BIRN) linking major neuroimaging research centres throughout the U.S. Built on prior projects led by his group at UCSD, BIRN was developed to establish a functional data integration environment in the context of research testbeds to help promote multi-disciplinary, multi-investigator collaborations and data sharing in biomedical research. These objectives were realised and developments are now embedded in many subsequently initiated large and small-scale programmes, including the Neurosciences Information Framework (NIF), the National Database for Autism Research (NDAR) and CAMERA.

In 1996, he founded the Centre for Research in Biological Systems (CRBS), an organised research unit at the University of California, San Diego. CRBS facilitates an interdisciplinary infrastructure in which people from biology, medicine, e-Science, engineering, mathematics and physics can work with those from computer science and information technologies.

In 1992, Professor Ellisman's research team introduced the idea of telemicroscopy and demonstrated network-enabled, remote use and sharing of the world's most powerful electron microscopes, including the 3MeV facility in Osaka, Japan.

In 1988, Professor Ellisman established the NIH National Centre for Microscopy and Imaging Research (NCMIR), an internationally acclaimed technology development centre and a widely used research resource that develops new technologies and provides researchers with access to many of the most advanced imaging technologies. NCMIR remains a very vital NIH-supported National centre, considered by many to be the international hot spot for biotechnology development related to advanced microscopic imaging, especially applied to the challenges to understanding presented by the brain.

After graduate studies in neurophysiology and behaviour, Professor Ellisman earned a Ph.D. in molecular, cellular and developmental biology from the University of Colorado, Boulder, studying with Keith R. Porter. He began his tenure as a Professor of Neurosciences and Bioengineering at UCSD in 1977. Since then he has received several UCSD teaching and lecturing awards including the Department of Neurosciences Award for Outstanding Teaching in 1987 and 1992, and the University Lecturer in Biomedicine in 2001. He also led for the University of California System the development of a 40-university consortium, the National Partnership for Advanced Computing Infrastructure (NPACI). The NPACI was supported by the National Science Foundation (NSF) and Professor Ellisman served as the interdisciplinary coordinator and led the Neuroscience activities for NPACI as well as the San Diego Supercomputer Centre (SDSC) from 1995 to 2004.

In addition to being a Founding Fellow of the American Institute of Biomedical Engineering, Professor Ellisman has received numerous awards including a Jacob Javits award from the NIH and the Creativity Award from the NSF. He has been appointed scientific advisor to numerous national and international organisations, and is frequently invited to lecture on neuroscience, scientific instrumentation and bioinformatics topics.

Most recently, Professor Ellisman served a 5-month stint as Director of the International Neuroinformatics Coordinating Facility (INCF), successfully shepherding this Stockholm-based organisation through its first external 5-year review. INCF involves activities in 15 countries, was established in 2005 and by the Global Science Forum of the Organisation for Economic Co-operation and Development (OECD) and fosters worldwide collaboration and data sharing in neuroscience to advance understanding of the human brain and its diseases. Professor Ellisman is now serving as the Director of Strategic Planning.

DR STUART FELDMAN

Dr Stuart Feldman is responsible for engineering activities at Google's offices in the eastern part of the Americas, with projects affecting most of the company's focus areas. He is also responsible for several important Google products.

Before joining Google, he worked at IBM for eleven years. Most recently, he was Vice President for Computer Science in IBM Research, where he drove the long-term and exploratory worldwide science strategy in computer science and related fields, led programmes for open collaborative research with universities, and influenced national and global computer science policy.

Prior to that, Dr Feldman served as Vice President for Internet Technology and was responsible for IBM strategies, standards, and policies relating to the future of the Internet, and managed a department that created experimental Internet-based applications. Earlier, he was the founding Director of IBM's Institute for Advanced Commerce, which was dedicated to creating intellectual leadership in e-commerce.

Before joining IBM in mid-1995, Dr Feldman was a computer science researcher at Bell Labs and a research manager at Bellcore. In addition he was the creator of Make as well as the architect for a large new line of software products at Bellcore.

Dr Feldman did his academic work in astrophysics and mathematics and earned his AB at Princeton and his Ph.D. at MIT. He is Past President of ACM (Association for Computing Machinery) and received the 2003 ACM Software System Award. He is also a Fellow of the IEEE, a Fellow of the ACM, a Fellow of the AAAS, a member of the Board of Directors of the AACSB (Association to Advance Collegiate Schools of Business, International). He serves on a number of government advisory committees.

PROFESSOR IAN FOSTER

Professor Ian Foster is Director of the Computation Institute, a joint institute of the University of Chicago and Argonne National Laboratory, where he is also the Arthur Holly Compton Distinguished Service Professor of Computer Science and an Argonne Distinguished Fellow. He received a BSc (Hons I) degree from the University of Canterbury, New Zealand, and a Ph.D. from Imperial College, United Kingdom, both in computer science. His research deals with distributed, parallel, and data-intensive computing technologies, and innovative applications of those technologies to scientific problems. Methods and software he has developed underpin many large national and international cyberinfrastructures.

Professor Foster is a fellow of the American Association for the Advancement of Science and the British Computer Society. His awards include the British Computer Society's award for technical innovation, the Global Information Infrastructure (GII) Next Generation award, the British Computer Society's Lovelace Medal, R&D Magazine's Innovator of the Year, and an honorary doctorate from the University of Canterbury. He was a co-founder of Univa UD, Inc., a company established to deliver high-quality grid and cloud computing solutions.

PROFESSOR ALBERT HECK

Professor Albert Heck is professor at Utrecht University, the Netherlands, and scientific director of the Netherlands Proteomics Centre and the Bijvoet Centre for Biomolecular Research. The general research theme of his group is to develop and implement innovative mass spectrometric methods for the more efficient and detailed characterisation of proteins in relation to their biological function. The emphasis is on the structural characterisation of proteins and post-translational modifications as well as the investigation of protein complexes and protein interactions. In short they apply protein mass spectrometry to problem in proteomics and in structural biology. As scientific director, Professor Heck and his group play a pivotal role in the Netherlands Proteomics Centre, which focuses on developing proteomics technologies, focused on protein expression quantification, membrane protein proteomics, post-translational modifications, protein networks, high-throughput protein analysis and biomarkers. The Heck laboratory has a track record in proteomics and especially in the analysis of protein post-translational modifications. They introduced TiO2 as enrichment material for the targeted analysis of phosphopeptides, and implemented this technique into a miniaturised on-line automatic system, and on a micro-chip device. They introduced the use of a protease named LysN, that in conjunction with ETD provides unique sequence ladders, that are straightforward to interpret and allow facile de novo sequencing and improve the analysis of protein phosphorylation. They also have an extensive track-record in quantitative proteomics, introducing metabolic stable isotope labelling in multicellular organisms such as Drosophila and C. elegans, using SILAC for studying stem cell and B cell differentiation, and stable isotope labelling by using chemical approaches. The latter method they most recently also implemented to follow differential pTyr phosphorylation in differentiating stem cells. Bioinformatics support and expertise to provide state-of-the-art data, pathway and network analysis is available.

The Heck laboratory is also a pioneer in macromolecular or native mass spectrometry, which enable the analysis of intact protein assemblies by mass spectrometry. The Heck group develops mass spectrometers dedicated for this work and applies these technologies to study the structure and dynamics of, for instance, transcription complexes and virus assembly. Professor Heck has published over 250 papers, and has been awarded several prizes, amongst them the Descartes-Huygens Prize and the Golden Medal of the Dutch Royal Society for E-Science.

PROFESSOR DIETER HEERMANN

Professor Dieter Heermann is Professor of Theoretical Physics at the University of Heidelberg. He is a Member of the Board of the Interdisciplinary Centre for Scientific Computing and Adjunct Professor at the Jackson Laboratory (Maine, USA). Beside serving on various committees at the University of Heidelberg, he serves for the European Community and national research organisations. He is an active member of various scientific organisations and has organised many international workshops and conferences.

He started his scientific career earning degrees in Informatics, Mathematics and Physics receiving his Ph.D. in Theoretical Physics from Boston University. One of his particular interests is the development and application of computational methods. He is well know for his books on Computational Physics.

The focus of his current research is biophysics with particular emphasis on the understanding of structure and function in the cell nucleus.

PROFESSOR JULIA LANE

Professor Julia Lane is the Programme Director of the Science of Science & Innovation Policy programme at the National Science Foundation. Her previous jobs included Senior Vice President and Director, Economics Department at NORC/University of Chicago, Director of the Employment Dynamics Program at the Urban Institute, Senior Research Fellow at the U.S. Census Bureau and Assistant, Associate and Full Professor at American University. Professor Lane has published over 60 articles in leading economics journals and authored or edited five books. She has been the recipient of over $25 million in grants and contracts; from foundations such as the National Science Foundation, the Sloan Foundation, the MacArthur Foundation, the Russell Sage Foundation, the National Institute of Health; from government agencies such as the Departments of Commerce, Labour, and Health and Human Services in the U.S., the ESRC in the U.K., and the Department of Labour and Statistics New Zealand in New Zealand, as well as from international organisations such as the World Bank. She has organised over 30 national and international conferences, received several national awards, given keynote speeches all over the world, and serves on a number of national and international advisory boards. She is one of the founders of the LEHD programme at the Census Bureau, which is the first large scale linked employer-employee dataset in the United States. A native of England who grew up in New Zealand, Professor Lane has worked in a variety of countries, including Australia, Germany, Malaysia, Madagascar, Mexico, Morocco, Namibia, Sweden and Tunisia. Her undergraduate degree was in Economics with a minor in Japanese from Massey University in New Zealand; her M.A. in Statistics and Ph.D. in Economics are from the University of Missouri in Columbia.

PROFESSOR LUCIANO MILANESI

Professor Luciano Milanesi received the BS degree in Atomic Physics in 1981. In 1986 he received the Ph.D. degree in Health and Hospital Physics, participating to the development of the project "Cyclotron facility for Positron Emission Tomography application in Nuclear Medicine".

Since 1987 he has been Staff scientist at the Italian National Research Council (CNR) and since 2004 he has been head of the Bioinformatics Division the Bioinformatics for the "Centre for Bio-molecular Interdisciplinary Studies and Industrial applications".

In 2007 he became Director of the CNR Interdepartmental Bioinformatics Research Network among 12 CNR research Institutes in Life science, Medicine and ICT.

He has been principle investigator for several European Projects: 1996 - 1999 European Commission "TRADAT" TRAnscription Databases and Analysis Tools, "O2I" Online Research Information Environment for the Life Sciences; 2002-04 European Commission ORIEL "an Online Research Information Environment for the

Life Sciences"; 2003-2004 NATO Science Programme "computer modelling, 5'-UTR, co-expressed genes, macrophages, Gene networks, Epstein-Barr Virus, cis elements recognition, B-DNA conformation"; 2004-06 European Commission INTAS "Modelling and analysis of mammalian cell-cycle regulatory networks in normal and pathological states by bio- and chemoinformatics". 2003-05 MIUR FIRB Post-Genomic "Bioinformatics for Genome and Proteome"; 2003-05 MIUR FIRB "GRID-IT: Enabling Platforms for high-performance computational GRIDS oriented to scalable virtual organisations."

He has been the coordinator of the European BIOINFOGRID project: "Bioinformatics Grid Applications for life science" and the MIR-FIRB "Laboratory of Bioinformatics Technologies" project. He is the principal investigator for the "Enabling Grids for E-Science EGEE II and EGEE III", pan-European Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) Projects.

Since 2002 Professor Milanesi has taught Bioinformatics for the course of Medical Biotechnology at the University of Milan, and since 2003 he has taught fundamentals in informatics at specialisation degrees for the University of Milan. He is a cofounder of BITS Bioinformatics Italian Association and cofounder of SYSBIOHEALTH System Biology for Health. He is a member of the Board of the International Neuroinformatics INCF Secretariat Karolinska Institutet, Sweden.

Professor Milanesi has published contributions in several books and scientific publications in Bioinformatics. He is the author of more than 260 publications in the field of Bioinformatics, Systems Biology and Medical Informatics.

DR JAYANTH GOPINATH PARAKI

Dr. Jayanth Gopinath Paraki is the Editor-in-Chief of the Handbook of Research in Knowledge Management in Telemedicine (2007-2010) with IGI-Global, Pennsylvania, USA. He is also the Founder of Omega Associates, India (2005) which is a registered partnership firm to promote Knowledge Management and Networking. He is currently engaged in research in Information Retrieval on the World Wide Web, Datamining and Biomedical Ontology. He practices Family Medicine in addition to his research work and studies.

His current focus is Foreign Direct Investment and exploring opportunities for collaborative projects in E-Science between developed and developing countries.

Between 2001- 2007 he has travelled to Sweden, Australia and USA presenting conference papers, key note address and building a platform for knowledge sharing and delivery processes. He supports use of ICT in Sustainable Development and Telemedicine. He also published a chapter International Institute for Knowledge Management in a book Information Technology and Economic Development which deals with UN Millenium Development Goals and role of technology in health education and E-Learning.

He completed his under-graduate medical education in 1983 and worked as a Senior Resident-General Surgery until 1992. It was while he was a Senior Resident at the Manipal Hospital, Bangalore that his interest in computers and databases was stimulated and he worked on MS-DOS while on emergency calls at night to understand the basics of data structure and management.

DR WOLFGANG VON RÜDEN

Dr. Wolfgang von Rüden holds a Ph.D. in physics from Mainz University, Germany. He came to CERN in 1975 and worked during the first part of his career on particale detectors and real-time data acquisition systems. From 1982-1989 he was project leader for the data acquisition system of the Aleph Detector at the LEP accelerator, one of the largest and most advanced systems in high energy physics at its time.

From 1990-1992, co-founder and president of IBEX Computing SA, a software company based in France, specialising in real-time systems and Computer Aided Software Engineering. He returned to CERN in 1992 where he introduced industrial control systems for beam lines and physics experiments.

From 1994 until 1998 he was Technical Director at GSI, a German National Research Institute in Heavy Ion Physics, with responsibility for most of the technical teams and the general infrastructure of the laboratory, including the computing services. He also coordinated the project "Tumour Therapy with Heavy Ions" at GSI during this time.

Back to CERN at the end of 1998, he joined the IT Department leading the Control System Group, then the Physics Data Processing Group from autumn 2000, before being nominated as Head of the IT Department (www.cern.ch/it) at the beginning of 2003. During his mandate most of the developments for LHC Computing took place, followed by deployment of the world-wide grid infrastructure. The CERN IT department also coordinated many EU-funded projects, notably EGEE.

Since January 2009, Dr von Rüden is Head of CERN openlab, a joined venture between CERN and leading IT companies (www.cern.ch/openlab). Dr von Rüden has served in various international reviews and he is invited regularly as referent to CIO-level events.

PROFESSOR ALEXANDER SZALAY

Professor Alexander Szalay is the Alumni Centennial Professor of Astronomy at the Johns Hopkins University. He is also Professor in the Department of Computer Science. He is a cosmologist, working on the statistical measures of the spatial distribution of galaxies and galaxy formation. He was born and educated in Hungary. After graduation he spent postdoctoral periods at UC Berkeley and the University of Chicago, before accepting a faculty position at Johns Hopkins. In 1990 he has been elected to the Hungarian Academy of Sciences as a Corresponding Member. He is the architect for the Science Archive of the Sloan Digital Sky Survey. He is Project Director of the NSF-funded National Virtual Observatory. He has written over 500 papers in various scientific journals, covering areas from theoretical cosmology to observational astronomy, spatial statistics and computer science. In 2003 he was elected as a Fellow of the American Academy of Arts and Sciences. In 2004 he received an Alexander Von Humboldt Award in Physical Sciences, in 2008 a Microsoft Award for Technical Computing. In 2008 he became Doctor Honoris Clausa of the Eötvös University, Budapest.

PROFESSOR PAUL TACKLEY

Professor Paul Tackley is chair of geophysical fluid dynamics in the Department of Earth Sciences at ETH Zürich. He studies the structure, dynamics and evolution of Earth and other terrestrial planets and moons as related to convective processes in the mantle, lithospheric dynamics and plate tectonics. His approach emphasises numerical simulation, using state-of-the-art numerical methods and high performance parallel supercomputers to obtain more realistic, three-dimensional numerical models of dynamical processes than previously possible. His group is presently involved in many different projects funded by the EU (TOPO-4D, Crystal2Plate) and SNF, covering lithosphere dynamics, mantle dynamics, planetary dynamics, planetary differentiation and extra-solar planets, and the development of appropriate numerical.

After obtaining his BSc at Cambridge University in 1987 and Ph.D. at Caltech in 1994, he ascended the professorial ladder in the Department of Earth and Space Sciences and Institute of Geophysics and Planetary Physics at the University of California Los Angeles, moving to ETH Zurich in 2005. He was the recipient of a Packard Foundation Fellowship and has served as Associate Editor for the Journal of Geophysical Research and Geoe-Science, Geophysics, Geosystems and is on the editorial board of Geophysical and Astrophysical Fluid Dynamics. He has made several academic visits to Japan sponsored by the Japan Society for Promotion of Science, Inoue Foundation, and recently as visiting Professor at the Earthquake Research Institute, University of Tokyo. He was chief organiser of two international scientific workshops in 2008 and 2009.

MR HAN WENSINK

Mr G. J. (Han) Wensink trained as a mathematician and has been working in the fields of Earth observation and operational oceanography for more than 24 years. He has managed large metocean consultancy and R&D projects. He is specialised in the modelling and processing of satellite data for many oceanographic, coastal and meteorological applications. His current responsibilities at ARGOSS include project acquisition, business development, and operational and financial management. At present, he is Vice Chairman of the European Association of Remote Sensing Companies (EARSC) and member of the board of the Geomatics Business Park.

PROFESSOR ANDERS YNNERMAN

Professor Anders Ynnerman received a Ph.D. in physics from Gothenburg University. During the early '90s he was doing research at Oxford University, UK, and Vanderbilt University, USA. In 1996 he started the Swedish National Graduate School in Scientific Computing, which he directed until 1999. From 1997 to 2002 he directed the Swedish National Supercomputer Centre and from 2002 to 2006 he directed the Swedish National Infrastructure for Computing (SNIC) and he is the chair of the strategic technical advisory committee for SNIC. Professor Ynnerman is currently a board member of the Swedish Research Council.

Since 1999 he has held a chair in scientific visualisation at Linköping University and in 2000 he founded the Norrköping Visualisation and Interaction Studio (NVIS). NVIS currently constitutes one of the main focal points for research and education in computer graphics and visualisation in the Nordic region. He is also one of the co-founders of the Centre for Medical Image Science and Visualisation (CMIV) and is serving as the chair of the scientific council for CMIV. Professor Ynnerman's current research interest lies in the area of visualisation of large scale and complex data sets with a focus on volume rendering and multi-modal interaction as well as fundamental computer graphics. He was awarded the Akzo Nobel Science award in 1997 for his extensive research contributions and in the same year he was awarded the Golden Mouse award for Swedish IT personality of the year. He is one of the founders of the Swedish e-Science Research Centre, which in 2009 was awarded a long term governmental grant to build e-Science communities in Sweden.

<table>
<tr><th>Date</th><th colspan="6">Sessions (Projects/Universities)</th></tr>
<tr><td>Sunday 6 December</td><td colspan="6">Afternoon welcome/briefing session</td></tr>
<tr><td rowspan="7">Monday 7 December</td><td colspan="6">Overview of research strategy</td></tr>
<tr><td colspan="2">Digital Curation Centre/ National e-Science Centre/ e-Science Institute (Edinburgh/Glasgow)</td><td colspan="2">Axiope/AstroGRID/ Bioinformatics Group (Edinburgh)</td><td colspan="2">Nano-CMOS/PolicyGrid II (NCeSS node)/DAMES (NCeSS node) (Aberdeen/Stirling/ Glasgow)</td></tr>
<tr><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Edinburgh University)</td><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Glasgow University)</td><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Aberdeen and Stirling)</td></tr>
<tr><td colspan="6">Open discussion and Q&A</td></tr>
<tr><td colspan="6">Poster session with early career academics</td></tr>
<tr><td colspan="6">Programme Leadership</td></tr>
<tr><td colspan="6">Private session for Panel</td></tr>
<tr><td rowspan="7">Tuesday 8 December</td><td colspan="6">Overview of research strategy</td></tr>
<tr><td colspan="2">North West e-Science Centre/MyGrid (Manchester)</td><td colspan="2">CLEF/CLEF-Services/ PsyGRID/iSPIDER (Manchester)</td><td colspan="2">Obesity e-Lab (NCeSS node)/e-Infrastructure Project (NCeSS node)/ NCeSS hub (Manchester)</td></tr>
<tr><td colspan="3">Biomathematics and Bioinformatics Group/eSTAR (Aberystwyth/Liverpool John Moores/ Rothamsted Research Institute)</td><td colspan="3">OMII-UK/LifeGuide (Southampton)</td></tr>
<tr><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Manchester)</td><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Southampton and Liverpool John Moores)</td><td colspan="2">Knowledge Exchange and Economic Benefit (Collaborators with Aberystwyth and Rothamsted)</td></tr>
<tr><td colspan="6">Open discussion and Q&A</td></tr>
<tr><td colspan="6">Poster session with early career academics</td></tr>
<tr><td colspan="6">Private session for Panel</td></tr>
<tr><td rowspan="3">Wednesday 9 December</td><td colspan="6">Overview of research strategy</td></tr>
<tr><td colspan="2">Intergrative Biology/ climateprediction.net/ OeSS (NCeSS node)/ OeRC (Oxford)</td><td colspan="2">VOTES/NeuroGRID/ Bioinformatics Group/ Technology and Documents (Oxford)</td><td colspan="2">CARMEN/BASIS (Newcastle)</td></tr>
<tr><td colspan="3">Knowledge Exchange and Economic Benefit (Collaborators with Oxford)</td><td colspan="3">Knowledge Exchange and Economic Benefit (Collaborators with Newcastle)</td></tr>
</table>

<table>
<tr><th>Date</th><th colspan="3">Sessions (Projects/Universities)</th></tr>
<tr><td rowspan="4">Wednesday 9 December (Continued)</td><td rowspan="3">Visit to Imperial College, London
MESSAGE/Discovery.net/ Centre for Bioinformatics/ GENeSIS (NCeSS node)/ RealityGrid/CATE (Imperial/UCL/Kings/ Natural History Museum)</td><td colspan="2">Open discussion and Q&A for remaining Panel</td></tr>
<tr><td colspan="2">Poster session with early career academics for remaining Panel</td></tr>
<tr><td>Visit to University of Reading
BioDiversity World/ Reading e-Science Centre/ GODIVA/GCEP/Storm tracking/GCOMS (Reading/Proudman Oceanographic Laboratory)</td><td>Visit to Rutherford Appleton Laboratory
NGS/NERC Data Grid/ Facilities/GridPP/ATLAS/ CMS (RAL/Glasgow/ Lancaster/Bristol)</td></tr>
<tr><td colspan="3">Private session for Panel</td></tr>
<tr><td rowspan="5">Thursday 10 December</td><td>Overview of research strategy (York/Leeds/Sheffield)</td><td>Overview of research strategy (Cambridge/Wellcome Trust Sanger Institute)</td><td>Overview of research strategy (Bristol/Nottingham)</td></tr>
<tr><td>White Rose Grid/ DAME/eHTPx (York/Leeds/Sheffield)</td><td>CancerGRID/eMinerals/ NIEeS/e-Family (Cambridge/Wellcome Trust Sanger Institute)</td><td>MiMeg (NCeSS node)/ e-Stat (NCeSS node)/ GENIE/DReSS (NCeSS node) (Bristol/Nottingham)</td></tr>
<tr><td>Knowledge Exchange and Economic Benefit (Collaborators with York/ Leeds/Sheffield)</td><td>Knowledge Exchange and Economic Benefit (Collaborators with Cambridge/Wellcome Trust Sanger Institute)</td><td>Knowledge Exchange and Economic Benefit (Collaborators with Bristol/Nottingham)</td></tr>
<tr><td colspan="3">Demonstrations: Virtual Vellum/Belfast e-Science Centre/climateprediction.net</td></tr>
<tr><td colspan="3">Poster session of STFC and EPSRC projects</td></tr>
<tr><td>Thursday 10 December (late pm)</td><td colspan="3">Commenced drafting report</td></tr>
<tr><td>Friday 11 December</td><td colspan="3">Agreed main findings and recommendations for future actions; presented these to Steering Committee</td></tr>
<tr><td>Saturday 12 December</td><td colspan="3">Panel departed</td></tr>
</table>

A range of supporting evidence and information was provided to the Review Panel both before and during the review. This included:

Overview: Funding of Science and Innovation in the UK: This describes the key developments that have taken place over the last 5 years concerning public funding arrangements for science and innovation in the UK and provides a high-level overview of how the science budget is secured and distributed. It also contains broad descriptions of various Research Council mechanisms for supporting research and training.

Background Data: This provided RCUK-related grant and studentship data, information about relevant TSB activities and information on other potential funders and collaborators. It was prepared as a companion to '*Funding of Science and Innovation in the UK*' to give more detailed contextual data relevant to e-Science research in the UK.

Consultation Responses: As part of the preparation for the review a public consultation was held to gather evidence for the Panel. Responses were specifically invited from the Institutions who visited the Panel, from others who did not visit the Panel, and from any other interested person/organisation via a public call on the Research Council web sites. A standard template was used to ensure that submissions addressed the Evidence Framework agreed by the Steering Committee. The Panel was provided with both a summary and the full text of all responses received. Responses were received from:

Institutions/individuals:

• Imperial College, London	John Darlington Michael Sternberg Yike Guo John Polak
• King's College London	Stuart Dunn
• Liverpool John Moores University	Iain Steele
• Natural History Museum, London	Malcolm Scoble
•Newcastle University	Paul Watson
• Rothamsted Research	Chris Rawlings
• Rutherford Appleton Laboratory (RAL), STFC	Neil Geddes
• The University of Manchester	Carole Goble John Brooke Simon Hubbard
• University of Aberdeen	Peter Edwards
• University of Cambridge	Mark Hayes
• University of Edinburgh	Richard Kenway
• University of Glasgow	Richard Sinnott
• University of Oxford	Ian Walmsley William Dutton
• University of Reading	Jon Blower Keith Haines Vassil Alexandrov David Spence Steve Gough Mike Roch
• University of Southampton	Philip Nelson
• University of Stirling	Paul Lambert
• White Rose Universities (Leeds, Sheffield and York)	Jie Xu

Other Stakeholders:

• Anatomical Society of Great Britain and Northern Ireland	Jonathan Bard
• Chair, Engineering Panel in RAE 2008	Dame Ann Dowling
• Diamond Light Source	Colin Nave
• Institute of Physics (IoP)	Tajinder Panesor
• Isaac Newton Institute for Mathematical Sciences, Cambridge	Sir David Wallace
• Linnean Society of London	Ruth Temple
• Panel Chair of last RAE exercise	Nigel Hitchin

- The UK Computing Research Committee (UKCRC) Muffy Calder

Research Group submissions: e-Science Research Groups at each institution with which the Panel met were requested to prepare in advance brief details of their strategic plans and individual research activities. A standard template was used to ensure that submissions addressed the Evidence Framework agreed by the Steering Committee. The following Institutions/Consortia submitted details relating to the projects/groups listed below:

- **Aberystwyth University**
  Computational Biology
- **Belfast e-Science Centre**
  Gridcast, OpenRiskGrid, Grid Enabled Distributed Data Mining (GEDDM), PRISM
- **Durham University**
  VirtU
- **Imperial College, London**
  AMUSe, CareGrid, Centre for Bioinformatics, Discovery Net, MESSAGE
- **King's College London**
  Arts and Humanities e-Science Support Centre
- **Lancaster University**
  ATLAS
- **Liverpool John Moores University**
  eSTAR
- **Natural History Museum**
  CATE
- **Newcastle University**
  CARMEN, BASIS, Culture Lab, Dynamic Virtual Organisations
- **Proudman Oceanographic Laboratory**
  GCOMS
- **Rothamsted Research**
  Bioinformatics
- **Rutherford Appleton Laboratory (RAL), STFC**
  Facilities, National Grid Service, NERC DataGrid
- **The University of Manchester**
  CLEF, e-Infrastructure, HyOntUse, iSPIDER, MyGrid, NCeSS, Obesity e-Lab, PsyGRID, ALMA
- **University College, London**
  CEDAR, Divergent Grid, e-Materials, GENeSIS, RealityGrid, E-Curator, Solar Dynamics Observatory
- **University of Aberdeen**
  PolicyGrid
- **University of Bristol**
  e-Stat, LHCb, MiMeG, CMS
- **University of Cambridge**
  eMinerals, Pervasive Debugging, Systems Research
- **University of Edinburgh**
  AstroGRID, Axiope, Bioinformatics Group, Digital Curation Centre, e-Science Institute, Managed Bandwidth, National e-Science Centre, Psychiatry, ReQueST, e-Science Research (joint with University of Glasgow)
- **University of Glasgow**
  AMUSe, Nano-CMOS, GridPP
- **University of Nottingham**
  DReSS
- **University of Oxford**
  Climateprediction.net, eDiaMoND, Integrative Biology, Bioinformatics Group, NeuroGrid, Oxford e-Research Centre, OeSS, Technology and Documents, VOTES
- **University of Reading**
  BiodiversityWorld, GCEP, Reading e-Science Centre, GODIVA
- **University of Southampton**
  LifeGuide, OMII-UK, PASOA
- **University of Stirling**
  DAMES
- **White Rose Universities** *(Universities of Leeds, Sheffield and York)*
  White Rose Grid, Virtual Vellum, g-Viz, e-HTPx, DAME

The following international and UK collaborating bodies made representatives available to the Panel during their meetings with the institutions:

- Argonne National Laboratory, USA
- Arjuna Technologies
- Beihang University, China
- CIS Informatics, Glasgow
- Demographic Users Group
- Eagle Genomics
- European Distributed Institute of Taxonomy (EDIT)
- Fujitsu Laboratories of Europe
- Goldsmiths College, London
- HP
- IDBS
- Louisiana State University, USA
- Macaulay Land Use Research Institute
- Microsoft Research
- o2
- Pfizer
- Rolls Royce
- Royal Botanic Gardens, Kew
- Royal Hospital for Sick Children, Glasgow
- Southern General Hospital, Glasgow
- Syngenta
- Syracuse University, USA
- Transport for London (TfL)
- University of Amsterdam

RESEARCH
COUNCILS UK

Research Councils UK
Polaris House
North Star Avenue
Swindon
SN2 1ET

Designed and printed by JRS Document Solutions