Issue 2 – 2007

HIGHLIGHTS

from the UK e-Science Programme

environment: e-Science delivers

health: acute inflammation and cholesterol

drug discovery: beating antibiotic resistance

Fostering collaboration to streamline research

Flood prediction, drugs against MRSA, remediation for environmental pollutants… these are some of the topics where the UK e-Science Programme is having an impact, as outlined in this second issue of *Highlights*.

By giving researchers access to resources held on widely-dispersed computers as though they were on their own desktops, e-Science enables research that would otherwise, in all probability, have been impossible. The resources include many digital data collections, very large scale computing resources, scientific instruments and high performance visualisation.

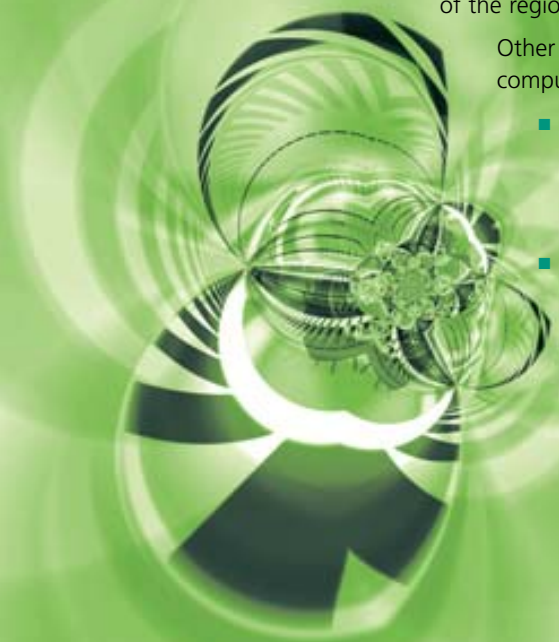
A grid allows these different resources to work together seamlessly across networks, enabling people to share them, often across traditional boundaries, and form virtual organisations. The vision is to facilitate collaborative working in multi-disciplinary teams by providing easy access to a grid via web interfaces, and powerful tools to organise computing tasks.

After six years of the UK e-Science Programme, this vision is coming closer to reality. The eMinerals project, for example, has developed an infrastructure that allows remote collaborators to share a complex task, such as elucidating all the interactions between a pollutant and soil, and immediately understand the results. What would previously have been too complex and time-consuming to be attempted can now be done within a few hours by researchers with little computing expertise.

eMinerals is one of four projects addressing pressing environmental issues that are featured in the middle pages of this issue. "e-Science is revolutionising the way environmental scientists are working," says Professor Robert Gurney in an introduction to the feature.

e-Science is also bringing about change by automating discovery processes.

- Drugs effective against antibiotic resistant bacteria have been discovered using a process that searches systematically through all licensed drugs for any that show action against the resistant bacteria.
- New insights into the anthrax bacillus have come to light thanks to a new e-Science method which systematically deduces from the bacterium's genome sequence all the proteins it secretes.
- A link between resistance to sleeping sickness in cattle and blood cholesterol levels was discovered because e-Science techniques enabled a thorough and systematic search of the region of the bovine genome associated with resistance.



Other applications illustrate how e-Science is enhancing the power of computer modelling.

- The GENIE*fy* project has run coupled intermediate-size models of the environment many times to get a result that large-scale models could not provide because they are too big to run often enough, even on a powerful supercomputer.
- Trans-continental grids, linking supercomputers, have enabled some of the largest molecular dynamics simulations ever undertaken, revealing new insights into the behaviour of clay sheets at certain scales.

The dedicated high speed optical networks used to link grids across continents are also being used to send data from widely-distributed radio telescopes to a central facility for immediate processing into an image, so revolutionising radio astronomy and the way in which radio astronomers collaborate.

You will find more information about these and other examples in the following pages, together with links to further information.

Acute inflammation and cholesterol: from African cows to Man

Research to track down the genetic root of resistance to sleeping sickness (trypanosomiasis) in African cattle has led to a discovery that could improve treatment for people in intensive care. Using techniques developed under the myGrid e-Science project, researchers from the Wellcome Trust host-pathogen project discovered that the raised blood cholesterol of trypanosomiasis-resistant cattle enables them to survive the high levels of inflammation caused by the disease. As patients in intensive care also experience acute inflammation as a result of serious injury or infection, could there be a similar link between cholesterol levels and survival chances in people?

Bioinformaticians working on the Wellcome project at the University of Manchester alerted Dr Paul Dark, a consultant in intensive care medicine, who undertook a study to investigate. Low cholesterol during the first few days in intensive care was found to be associated with non-survival. "We need to know the mechanisms behind this result before we can say whether it would be helpful to artificially raise cholesterol in people with acute inflammation," says Dr Dark.

Professor Andy Brass, who led the Wellcome project's bioinformatics team, praises the role of e-Science in making the sleeping sickness discovery. By enabling the function of all genes loosely associated with trypanosomiasis resistance to be thoroughly investigated, e-Science allowed researchers to avoid the pitfall of selecting a wrong hypothesis based on previous knowledge. "e-Science is incredibly powerful because it enables you to explore the full range of possible hypotheses. It allows you to go through all the genes in the regions where there's a perturbed pathway and see what's happening. The cholesterol pathway dropped out and e-Science took us to a place we would never have got to otherwise," he says.

www.nibhi.org.uk/research.aspx

New weapons against hospital superbugs

Three drugs that are effective against antibiotic-resistant superbugs, including MRSA, have come to light thanks to e-Science. They are the first antibiotics employing a truly novel mode of action to be discovered for decades. e-Therapeutics, a spin-out company from Newcastle University used grid computing and e-Science techniques to trawl through the portfolio of existing, licensed drugs for any that showed action against the superbugs.

"With these techniques we can search through all the data, warts and all, very rapidly and identify candidates within a fraction of the time it would take using conventional drug discovery methods," says Professor Malcolm Young, chairman of e-Therapeutics. The eXSys e-Science project demonstrated the reliability of the new method by showing that it could predict accurately the action and side effects of all 103 previously known antibiotics and many other drugs.

The techniques are also being taken up in the search for new drugs from the Amazon rain forest. e-Therapeutics has formed relationships with pharmaceutical companies in north-east Brazil to test substances extracted from rain forest plants for their efficacy against a range of diseases. The techniques reduce the time taken to analyse a substance from, typically, two years to two months.

"Drug discovery is a search problem. Conventional drug discovery is running out of steam because the search methods employed to date are so inefficient. What we're trying to do is industrialise drug discovery by implementing all these processes in computers, powered by e-Science," says Professor Young.

www.neresc.ac.uk/projects/eXSys/

SimCity for real promises virtual trials of social policy

Social policymakers and town planners will soon be able to play 'SimCity' for real using grid computing and e-Science techniques to test the consequences of their policies on a real, but anonymous, model of the UK population. The researchers at the University of Leeds are using data recorded at the 2001 census to build a model of the whole UK population, but with personal details omitted so no individual or household can be identified. "We're building a core model which represents the whole of the UK at the level of (synthetic) individuals and households with many attributes and behaviours," says Dr Mark Birkin who is leading the project.

Data about these attributes – such as car ownership, house prices and use of health, education, transport and leisure facilities – are held by different agencies in different locations and often in different formats. "Historically, people have assembled data on a single PC or workstation. e-Science provides exciting opportunities to access multiple databases from remote, virtual locations, making it possible to develop highly generic simulation models which are easy to update," says Dr Birkin.

The model can be projected into the future to explore the effect of different demographic trends and also to test the consequences of policy decisions.

www.ncess.ac.uk/research/nodes/MoSeS/

e-Science faces up to environmental challenges

With the global issues of climate change and increased use of natural resources never far from the headlines, environmental scientists face a critical challenge of predicting the future state of the environment and identifying practical solutions. e-Science is developing new tools and techniques that are revolutionising both the way environmental scientists are working, and the way they are working together to address these challenges.

Accurate prediction of the future state of the environment depends on the development of computer-based models to simulate how the environment will respond. To build these models we need: an understanding of the dynamics of environmental processes; high quality data sets to validate the models; the technical infrastructure to run the models and to manage increasing volumes of data collected; the ability to run the models many times in 'ensembles' to quantify the uncertainty associated with such predictions; and the ability to visualise and analyse the results. e-Science is providing many of the tools that will enable us to do these things more effectively.

The Natural Environment Research Council (NERC) e-Science Programme has focused on the development and application of new tools and techniques to undertake innovative environmental science. Results from two major NERC e-Science projects, GENIE*fy* and eMinerals, are highlighted in this issue, together with other e-Science work aimed at helping us to understand or manage our response to the environment.

The challenge for the future is to ensure that advances are built upon and e-Science becomes part of the normal way in which environmental scientists work.

Professor Robert Gurney, NERC Environmental Systems Science Centre.

A thorough approach to pollution remediation

A finding from the eMinerals project could help to avoid arsenic contamination of drinking water extracted from artificially-dug wells. Arsenic often appears in minerals rich in iron and sulphur, such as pyrite (fools' gold). Using e-Science techniques and grid computing, the project has found out precisely how arsenic is taken up and held in the pyrite structure and the factors likely to lead to its release.

"We now know that arsenic replaces the sulphur in pyrite rather than the iron, and that pyrite is likely to dissolve more easily when arsenic is present," says Dr Kate Wright who worked on the project. Further work could identify ways of stabilising arsenic-containing iron sulphide rock by introducing additives that slow the rate at which it dissolves.

Another finding could lead to improved methods of removing the now-banned industrial chemical, dioxin, from soil. The eMinerals project found that a dioxin molecule will bind more strongly to clay surfaces the more chlorine atoms it contains, irrespective of the position of the chlorine atoms in the dioxin molecule. Binding is also stronger the greater the electrical charge on the surface. However, water competes with dioxin to bind to surfaces and, in practice, a dioxin molecule's ability to bind to a surface is a balance between the binding strength of the dioxin to the surface, the water to the surface, and the dioxin to the water.

Both findings involved numerous simulations of the interactions between the different minerals in soil and rock with all the known variants of the contaminants. The project developed a grid and e-Science infrastructure to enable these simulations to be performed simultaneously within a few hours. Without this infrastructure, the task would have been impracticable.

www.eminerals.org

What keeps the global ocean circulation on – or off?

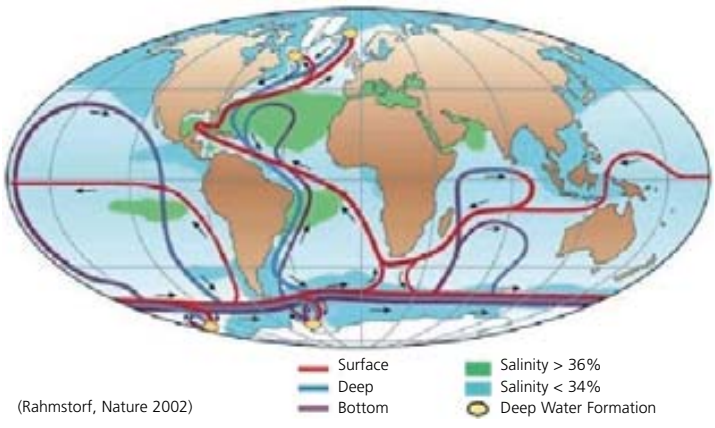
Could the melting Greenland ice sheet introduce enough freshwater into the north Atlantic to disrupt the salt and temperature gradients that drive the global ocean circulation? If global warming caused this to happen, the climate of the north Atlantic would, paradoxically, get colder.

The GENIE*fy* project is helping to provide some answers. Already, it has revealed new insights into the factors that tend to keep this so-called thermohaline circulation switched either on or off.

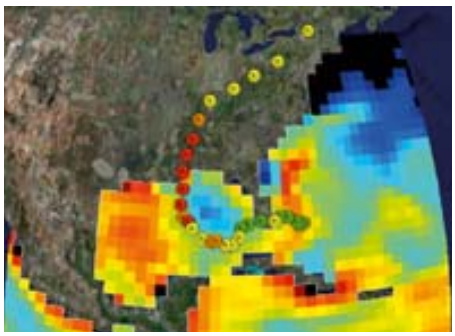
"This study has shown that when the circulation is in one state, the climate system acts to keep it that way. When the circulation switches off, the ocean and the atmosphere bring fresh water into the Atlantic, effectively exporting salt from the region and keeping the circulation turned off. When the circulation is on, it brings salt in, making the water denser and effectively keeping the circulation going," says Professor Tim Lenton from the University of East Anglia who heads the GENIE*fy* project.

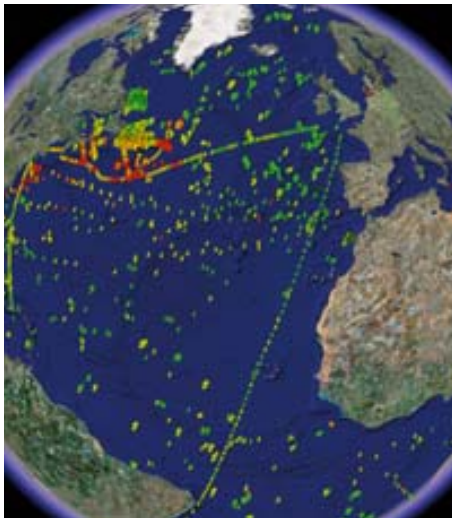
The project team came to these conclusions using grid computing and e-Science techniques to run many different computer models of different aspects of the environment at the same time. The models were coupled together and run on the distributed computing resources of the National Grid Service, enabling the discoveries to be made within three months instead of the five years the study would have taken using one computer alone.

www.genie.ac.uk/GENIEfy/index.htm



(Rahmstorf, Nature 2002)





Virtual globes visualise the Earth's environment

Google Earth is a computer-based 3D representation of the Earth on which you can superimpose your own information. Based on satellite imagery, it can also show you, rather alarmingly, an image of your house from space. Now, however, environmental scientists are using Google Earth and other similar virtual globes to visualise complex scientific data in new and revealing ways.

The image above left, for example, visualises two separate datasets simultaneously to show how the passage of Hurricane Katrina affected sea surface temperature in the Gulf of Mexico. One dataset shows sea surface temperature, with red colours being warmer. The other tracks the passage and intensity of the hurricane, with red dots representing greater intensity. "The image shows that the hurricane caused the sea on the right-hand side of the storm to cool which is where the strong winds would have caused upwelling of colder subsurface water," says Dr Jon Blower from the Reading e-Science Centre who is developing these techniques.

Google Earth is also helping environmental scientists to test their models against real data. The image below left shows at a glance whether a model of the ocean is good at predicting what is actually observed. Green pushpins indicate that the model's predictions of temperature and salinity accord well with measurements recorded by ships and buoys. The red pins, which are mainly located in the turbulent Gulf Stream, are not such a good match. "A virtual globe allows the scientist to view the data at a huge range of scales. Sometimes an overview is appropriate and sometimes the scientist needs to zoom in to a region where observations are very closely spaced," says Dr Blower.

In another application, scientists are displaying real-time data in Google Earth to help them direct scientific missions. The movement of penguins, recorded in real-time by scientists in the field, for example, is being combined with satellite images of nutrient concentrations to plot where the penguins, and hence the scientists, should be heading next.

www.resc.rdg.ac.uk/projects.php

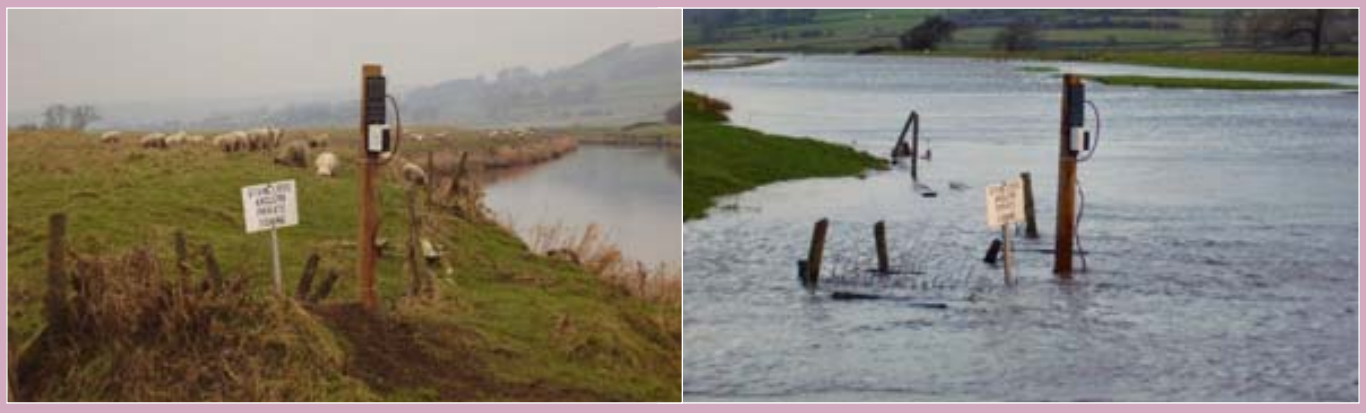
Real-time flood monitoring

An intelligent flood monitoring system that could give advance warning of the type of local floods that have been engulfing parts of the UK this summer is under test in the Yorkshire Dales. Unlike most current systems which issue general flood warnings over large areas, the new system promises rapid, low-cost warnings specific to particular flood-prone sites. Funded by the North-West Development Agency, it is based on a network of intelligent sensors which communicate with each other to form a computing grid using software developed under the Open Overlays e-Science project.

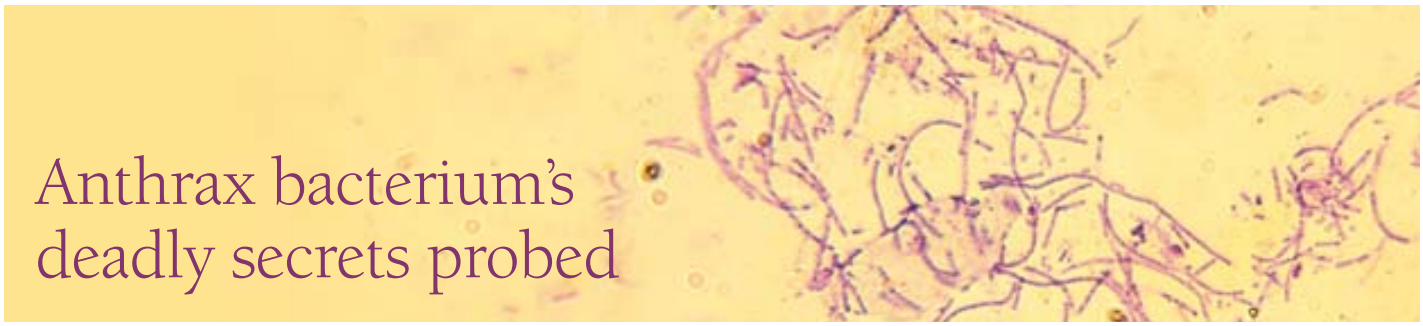
The sensors are placed in different locations across a flood plain to record water depth. A digital camera, placed on the river bank, monitors flow rate from the speed of flotsam between two points. Each sensor incorporates a powerful computer, no bigger than a packet of gum, which communicates wirelessly with other sensors in the network to form a computing grid. "As soon as the sensors detect water coming down the valley, the network gears up," says Danny Hughes, a member of the team developing the system at Lancaster University.

In order to provide flood warnings, the system makes use of flood forecasting models which can be run on the sensor computing grid and adjusted so that their predictions stay in line with what the sensors are recording. "An interesting possibility is to use such a local warning system to give advanced warning, even in catchments where the response to rainfall is very fast, making flood forecasting very difficult," suggests Professor Keith Beven of Lancaster who is also involved in the project.

www.comp.lancs.ac.uk/~hughesdr/nwgrid/



A depth monitoring sensor before and during a flood



Anthrax bacterium's deadly secrets probed

e-Science techniques have revealed how the proteins secreted by the anthrax bacterium equip it for its unusual lifestyle.

In many ways, *Bacillus anthracis* resembles a soil-growing organism. However, it lies dormant in soil, for hundreds of years in some cases, and becomes active only after it has been ingested by a suitable animal host. Starting only with a knowledge of the bacterium's genes, researchers at the North East Regional e-Science Centre have deduced that *Bacillus anthracis* secretes no proteins that would enable it to take up nutrients from soil, but plenty that enable it to grow in an animal host.

Using Taverna workflow technology developed under the myGrid project, they have developed a technique that, by a process of selection and elimination, deduces and characterises the proteins a bacterium secretes from knowledge of its genome sequence.

They tested their method on 12 members of the *Bacillus* family which range from the friendly *Bacillus subtilis* to the deadly *Bacillus anthracis*. The predicted secreted proteins from *Bacillus anthracis* help to explain its inability to grow in soil. "We're beginning to understand why *Bacillus anthracis* behaves in the way that it does – and how it has adapted only to grow in the host and not in the soil," says Professor Colin Harwood.

The researchers found a group of proteins of unknown function that are secreted only by pathogenic members of the *Bacillus* family. These proteins could be markers for pathogenicity in other bacteria. A website is under development to guide users through the process for any bacterium whose genome is known.

This work was co-funded by Nonlinear Dynamics.

www.neresc.ac.uk/projects/Microbase/

Three linked grids reveal unsuspected property of clay

The pooled resources of supercomputers on three computing grids across two continents have performed some of the largest computer simulations ever attempted to reveal a previously unsuspected property of clay.

The simulations showed that, at a certain size, microscopic sheets of clay start to undulate. This discovery could shed light on the properties of an important new class of materials, clay polymer nanocomposites, which are under investigation for a number of applications ranging from car bodies and other automotive uses, through oilfield technology to drinks packaging.

Computer models of clay sheets of differing size and complexity were run concurrently on supercomputers on the UK National Grid Service, the US TeraGrid and DEISA (EU Distributed European Infrastructure for Supercomputing Applications). In the largest model, the motions of nearly ten million atoms were taken into account. Grid middleware developed under the RealityGrid e-Science project enabled the researchers to access the supercomputers.

"As we moved from smaller to larger models we began to see collective undulations – the clay platelet sheets fluctuate up and down," says Professor Peter Coveney from University College London. This property, which was not known before in clay materials, has implications for the properties of clay on an ordinary scale which can be computed and then compared with experiment.

www.realitygrid.org

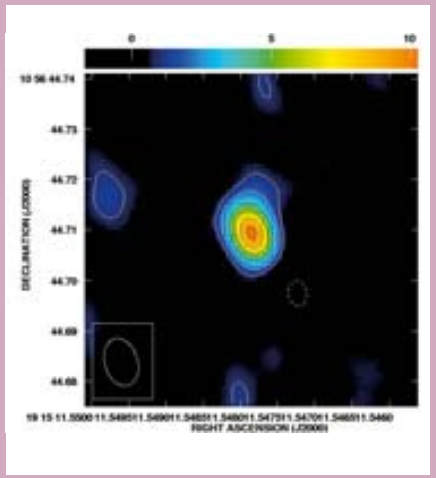
Radio astronomy gets responsive

Six radio telescopes across Europe have been transformed into the equivalent of one giant telescope thanks to the use of high speed optical networks, such as UKLight, the UK's optical network for research. Each telescope records data of the same object at the same time and transfers them via the networks to a central facility for immediate processing into an image. This real-time data transfer is now replacing road transport of data recorded on discs which often resulted in a wait of weeks or months for a final image.

Combining data in real-time produces images of unprecedented detail and sensitivity and allows radio astronomers to make adjustments quickly when image quality is poor. It also enables them to respond to rapidly varying phenomena.

The image (right) shows the detection of a flare associated with a binary star system in our galaxy, which probably contains a black hole or neutron star. "This happens on a time scale of about ten hours and recording data on disc could only image the object weeks after the event," says Anthony Rushton, an astronomer at the Jodrell Bank radio telescope at the University of Manchester. The ESLEA radio astronomy project developed the use of UKLight to transfer data between Jodrell Bank and the Dwingeloo data processing centre in the Netherlands.

www.eslea.uklight.ac.uk/sp_radio_astronomy_summary.html



e-Science becomes ‘business as usual’

Many of the techniques developed during the first six years of the UK e-Science Programme are now maturing sufficiently to be incorporated into standard research practice.

The future focus of the Core e-Science Programme, which is now funded by the Engineering and Physical Sciences Research Council (EPSRC), will be on the support of the generic facilities that users and potential users of e-Science tools and techniques need to further their research. These include, for example, the National Grid Service, which provides access to computational and data resources at a growing number of university sites throughout the UK; the Open Middleware Infrastructure Institute, which re-engineers software developed under the e-Science Programme to make it widely useable; and e-Science Centres which have traditionally provided support to researchers in the use of e-Science techniques.

The Core Programme will also provide support for networking activities for the community and to promote UK e-Science internationally.

The Research Councils[1], through the Core Programme, are working with the Joint Information Systems Committee (JISC)[2] to contribute to a national e-infrastructure that meets the requirements of the Government’s Investment Framework for Science and Innovation 2004-2014[3]. The e-infrastructure consists of the networks, services and software needed to do e-Science. JISC provides ICT infrastructure and services to UK further and higher education and research. In particular, it sets strategy and provides funding for the JANET high bandwidth academic network which forms the backbone for the UK e-infrastructure.

UK e-Science from 2001-06

The UK e-Science Programme, which began in 2001 as a five-year coordinated initiative, consisted of an e-Science Core Programme, managed by the EPSRC to serve the communities of all the Research Councils[1], and e-Science Programmes run by each Research Council. The former Department of Trade and Industry also ran its own e-Science Programme as part of the coordinated initiative.

The Core Programme supported the development of generic technologies, such as the software known as middleware that is needed to enable very different resources to work together seamlessly across networks and create computing grids. It also established the National e-Science Centre and Institute and regional e-Science centres to engage and support a wide spectrum of academic, industrial and commercial users in e-Science projects. The e-Science programmes of the individual Research Councils funded projects to develop e-Science techniques and demonstrate their use across a broad range of research and in a variety of applications. The UK e-Science Programme as a whole funded more than 100 projects during these first five years.

[1] Arts and Humanities Research Council (AHRC) **www.ahrc.ac.uk**
Biotechnology and Biological Sciences Research Council (BBSRC) **www.bbsrc.ac.uk**
Economic and Social Research Council (ESRC) **www.esrc.ac.uk**
Engineering and Physical Sciences Research Council (EPSRC) **www.epsrc.ac.uk**
Medical Research Council (MRC) **www.mrc.ac.uk**
Natural Environment Research Council (NERC) **www.nerc.ac.uk**
Science and Technology Facilities Council (STFC) **www.stfc.ac.uk**

[2] Joint Information Systems Committee (JISC) **www.jisc.ac.uk**

[3] **www.hmtreasury.gov.uk/spending_review/spend_sr04/associated_documents/ spending_sr04_science.cfm**

Further information about the UK e-Science Programme is available at **www.rcuk.ac.uk/escience** and **www.nesc.ac.uk**

e-Science Core Programme Contacts

Envoy – Malcolm Atkinson **mpa@nesc.ac.uk**

Head of High End Computing and
Core e-Science Programmes – Jane Nicholson **jane.nicholson@epsrc.ac.uk**

Research Council e-Science Contacts

AHRC – David Robey **d.j.b.robey@reading.ac.uk**

BBSRC – Jef Grainger **jef.grainger@bbsrc.ac.uk**

ESRC – Suzanne Tanner **suzanne.tanner@esrc.ac.uk**

EPSRC – Jane Nicholson **jane.nicholson@epsrc.ac.uk**

MRC – Lelia O'Connell **lelia.oconnell@headoffice.mrc.ac.uk**

NERC – Ned Garnett **nedg@wpo.nerc.ac.uk**

STFC – Trish Mullins **trish.mullins@stfc.ac.uk**

Written by Judy Redfearn **judy.redfearn@epsrc.ac.uk**