eScience curricula at two Australian universities

Henry Gardner[1] **Chris Johnson**[1] **Geoff Leach**[2] **Pascal Vuylsteker**[1]

[1]Department of Computer Science, FEIT
Australian National University
Canberra ACT 0200, Australia
Email: henry.gardner@anu.edu.au
[2]Department of Computer Science,
RMIT University,
GPO Box 2476V, Melbourne, Victoria 3001,
Email: gl@cs.rmit.edu.au

Abstract

Two Australian computer science departments have worked together to introduce the world's first degrees in "eScience". In both universities, the eScience programs have had a positive, lasting impact on the computer science curricula. In the case of the Australian National University, the eScience programs have become self-funding and have helped to position the university well for future strategic initiatives in eScience and "e-Research". At RMIT although the eScience program itself did not prove to be viable enough to continue in its own right, courses introduced as part of the eScience graduate diploma program have seen strong enrolments and continue to be offered to students undertaking other programs.

Keywords: eScience, e-Science, e-Research, graduate, coursework

1 Introduction

"eScience", "e-Science", and "e-Research" are three new buzz-words which describe computer, and computational, science activities associated with high performance computing and networking, and the increasing role that they play in scientific research. The idea is that some "grand-challenge" activities in the sciences and humanities can be significantly enhanced by harnessing, developing and networking advanced-computing infrastructure, experimental facilities and databases. More generally, eScience captures the notion that computers, networks and the web have become fundamental tools in many areas of science and are actively researched in pursuit of achieving scientific aims. In some overseas countries, eScience activities have been associated with substantial funding initiatives. In Australia, the ARC is piloting a new e-Research support through its Special Research Initiatives program(Australian Research Council 2004).

This paper describes the experience of two Australian universities in developing the world's first education programs in eScience.

2 eScience in Australia and the world

In 1999, the Australian government Department of Education, Science and Training, DEST (then DETYA) established a program designed to foster linkages between academia and industry with the idea of attracting students into science-related areas which would then benefit emerging industries. This program, the Science Lectureships Initiative (SLI), had three objectives

1. to assist universities to develop courses to meet the needs of established and emerging industries;
2. to attract students into science-related courses to meet the needs of industry through the development of innovative approaches to course design and delivery; and
3. to encourage industry to invest in its own future through the sponsorship of education and training.

Three of the authors were members of a small team which put together a successful SLI bid. In what we thought was a forward thinking, although somewhat outrageous, use of language, we called it "eScience: Applications and Technology". Our proposal was valued at about $1 million over 3 years across two universities. The bid was awarded to us in 2000 with funding commencing in the later half of that year. Graduate education programs were introduced from 2001. By the end of the program, in July 2003, there were some 70 students at both universities taking Graduate Diploma and Masters courses in Information Technology (eScience).

Independently of our initiatives in education, a UK initiative in "e-Science" was being set up to develop new research infrastructure. In November 2000, the Director General of Research Councils in the UK, Dr John Taylor, announced 98 million pounds in funding for a new UK e-Science program. Funding was made available across the UK research councils including physical and engineering sciences, economics and social sciences and biomedical and medical sciences. The initiative had several objectives: to "grid-enable" experimental facilities; to purchase a new Teraflop-scale HPC system; to develop and broker generic technology solutions and generic middleware to enable e-Science and form the basis for new commercial e-business software; to support e-Science pilot projects and so on. The UK initiative has since been refunded and is now valued at A$450 million over 5 years(Australian Research Council 2004). Exemplar e-Science projects include the following(Research Councils of the UK 2004)

- AstroGrid - enabling the Virtual Observatory
- Axiope - Data Management and Data Sharing for Scientists
- BioSimGrid - making large biomolecular simulations accessible
- Climate Prediction.Net - addressing Climatic uncertainty
- Comb-e-Chem - exploiting combinatorial Chemistry
- DAME - diagnostics and decision support across the Grid
- Discovery Net - Information from Hight-throughput devices
- e-Diamond - state of the art technology for breast screening
- e-Family - access to protein sequences and structures
- GENIE - integrating earth models to give a complete picture
- GridPP - a Grid for Particle Physics
- IXI - Medical imaging on the Grid

It is worth illustrating the ideas of eScience with reference to the UK program because they are compatible with our original vision of the discipline. They have also been influential with the Australian Research Council in its development of a new initiative in "e-Research"(Australian Research Council 2004).

3 Early days in eScience education

After receiving the SLI grant, we quickly needed to define programs which were true to our ideas of "eScience", which were appealing to industry and which were able to be delivered practically. In our proposal, we had pledged that "the eScience: Applications and Technology project will develop a graduated set of courses for science graduates to be trained in contemporary Information Technology, to become practitioners of what we term eScience. The project will develop special-purpose skills and knowledge of science graduates, differentiating it from conventional graduate IT courses." So one of our goals was to run conversion graduate courses which were particularly geared to graduates in the sciences and engineering. We wanted our courses to teach them IT skills which we felt were particularly "eSciencey" but we also needed to produce graduates from our programs who were attractive for industry and, ideally, we would have liked to have industry involvement in our course design as well.

In the end, domestic considerations were very strong influences in our choice of curricula. At ANU, we had previous experience in introducing courses in the eScience-allied area of "computational science and engineering" (which has resulted in an undergraduate Bachelors of Computational Science degree). We had also developed infrastructure for scientific visualisation and virtual reality – a interactive, stereoscopic theatre (the "Wedge") which we had built in-house. The department had a new Bachelors program in Software Engineering which was developing strong linkages with industry. Elsewhere on campus, there was a popular, graduate course in Graphical Data Analysis which dealt with the statistical fundamentals of scientific visualisation. We united these influences into a Graduate Diploma of IT (eScience) which had the following core components

- Introduction to programming in an OO language
- Advanced programming using an eScience case study
- Software Analysis and Design
- Computer Graphics
- Graphical Data Analysis
- Internet, intranet and document systems

The prerequisites for this program are that students should have completed a previous degree in science (or engineering) with reasonable grades as well as some university level mathematics. Students need not have had any formal computing background but they are meant to be able to have good study skills and a "sciencey" orientation. They start their studies with an Introductory Programming course which starts before regular semester and which aims to deliver programming skills and experience equivalent to an advanced-level first year programming course. Shortly afterwards, they also take the advanced programming course which is built around an eScience case study theme. This course teaches quite advanced skills and concepts in object-oriented programming and beginning students find it difficult. But the combination of the two courses is meant to bring the students up to a second-year programming expertise by the end of their first semester. In the second semester, students then take Computer Graphics which has a programming level equivalent to a 3rd or 4th year course. They also begin work on their final project course which is meant to synthesise their coursework in a project which has an eScience orientation. The Graduate Diploma ran for the first time in 2001 and attracted 10 students with backgrounds in physics, chemistry, biology, mathematics, medicine (and one from an IT background). Some of these students completed in that year and all of their projects could be labelled as eScience (or e-Research).

At RMIT, a set of postgraduate coursework programs was already on offer, including "Software Development" and "Internet and Web Computing". The eScience Graduate Diploma needed to be differentiated from these existing offerings and it was strongly pitched towards people with a science or engineering background who were interested in topics which took specific advantage of that background. The RMIT program was designed to offer graduates the ability to pursue specialist scientific niches or to migrate into the general IT industry. One of the main ideas was that the web started in science laboratories and was continuing to be a major influence in the access and dissemination of scientific and technical material and information - including for example, the area of bioinformatics. Core components in RMIT the eScience Graduate Diploma were

- Java for Programmers

- Web Servers and Web Technology
- Graphics Programming using Java3D
- eScience Case Studies

The eScience Case Studies course fulfilled a similar role to the project course at ANU - it was a course in which students both examined a number of eScience case studies and where they undertook a project to bring together the material they had learnt through the program.

A range of electives were offered, including Web3D, Computational Biology, Computational Science and more traditional IT courses such as Software Engineering and Databases.

4 Program evolution

At ANU, a conversion Masters in Information Technology (eScience) was introduced in 2002. Entry to this program differed from the Grad Dip in that programming experience was required (to a level equivalent to one semester of second-year). The Masters ran for 18 months full-time as against 12 months for the Grad Dip. It shared 4 courses with the Grad Dip and also had a new core course in "Internetworked Virtual Reality" as well as a large number of "constrained elective" courses in areas such as High Performance Scientific Computing, Algorithms, Computer Networks, Teleoperation and Mobile Computing and Human Computer Interaction. The Masters project was a larger course than for the Grad Dip and had research aspirations.

The ANU programs underwent considerable evolution from 2001 to 2004 due to several factors: feedback from students; feedback from industry and from financial considerations. The direction of these changes can be summarised as

- Students wanted more flexibility in the course structure; they wanted courses in databases; some preferred to choose a programming-based course over the statistics course; some preferred to take project courses which were shorter and more applied.
- The detailed structure of particular eScience courses were strongly influenced by student feedback.
- Industry feedback. Industry representatives on the management committee emphasised the importance of Human Computer Interaction in the program structure.
- Once the DEST grant had finished, economics dictated that some courses had to be merged, others delivered less frequently and others discontinued (but, overall, these changes were minor).

The most dramatic change in the ANU course has been the composition of the student body. In 2003 there was a large influx of overseas students, particularly from the Indian subcontinent and from East Asia. This has accompanied a diversification of backgrounds and interests of the students: at the time of writing the ANU program has about 60 students of which about 70% are overseas students and about 45% have prior backgrounds in Information Technology (rather than in straight science and engineering).

At RMIT, the eScience Graduate Diploma started with 10 students in the first year, but thereafter struggled to gain enough extra students to become self-sustaining as a separate program. Several courses specifically funded by the eScience grant were offered to students from other programs and were quite popular with enrolments between 50-100 students (Web3D and Java3D in particular) and continue to be offered.

5 Examples of courses

Our eScience graduate programs are conversion degrees. Students are meant to come into them having a prior qualification in science and engineering and leave with an IT qualification which is competitive with our straight Bachelors graduates (3rd year in the case of the Grad Dip; 4th year in the case of the Masters). This being the case, some of the courses on offer are designed to be "transition courses" which aim to ramp up a student's ability to the level of a later-year IT undergraduate. Other courses are designed to be "specialist courses" which feature an aspect of what could be called eScience. A third set of courses are aliased to senior-level undergraduate and Honours courses in order to teach relevant, advanced computer-science material and in order for the students to demonstrate that they can compete with the IT undergraduates at these levels.

The following examples are some of the "eSciencey" courses which students can take as part of our eScience programs:

Networked Scientific Data Analysis and Presentation: This course makes use of an eScience case study to teach advanced programming techniques with a particular emphasis on object-oriented design patterns. The case study chosen is a networked data visualiser used in fusion energy research. The full software is Globus-enabled and is used on Grid architectures. Students build a "browser" for the waveform data which is archived in a special database of experimental results. They are then taught to "refactor" the software using design patterns with particular emphasis on the encapsulation and reuse of sub-systems. The course has a "transition" role for Masters students but is an advanced course for Grad Dip students.

Internetworked Virtual Environments: This course emphasises the design and implementation of real-time, visual simulation systems for distributed virtual environments. There are three main areas of concern: virtual reality, groupware and networking for distributed virtual environments. Students engage with the research literature and, in their project work, build a networked application for the Wedge virtual reality theatre at ANU.

HCI and Usability Engineering: This course provides an introduction to the field of Human Computer Interaction and introduces students to behavioural research methods and techniques used in usability testing. Some of these techniques involve a consideration of elementary statistics. The course gives students the essential theoretical background to approaches, methods and techniques followed by practical experience in conducting usability studies for interactive systems, with a special focus on 3D and Virtual Reality applications. Aspects of professional practice in integrating usability testing with the software development cycle are also considered. This course uses the SPLUS language for statistical analysis; students wishing to deepen their knowledge of statistical, experimental techniques are advised to take the Graphical Data Analysis course. This course was developed collaboratively between ANU and RMIT.

Graphical Data Analysis: Introduces the principles of data representation, summarisation and presentation with particular emphasis on the use of graphics. The course uses the SPLUS Language in a modern computing environment. Topics to be discussed include: Data representation (examples of good and bad graphics; principles of graphic construction; some pitfalls to be avoided; presentation graphics), Graphics environments (interactive graphics; windows; linked windows; graphics objects), Statistical graphics (stem and leaf plots, box plots, histograms; quantile-quantile faces; dynamic graphics including data rotation and brushing), Relationships between variables (smoothing scatterplots; simple regression; modelling and diagnostic plots; exploring surfaces; contour plots and perspective plots; multiple regression; relationships in time and space; time series modelling and diagnostic plots).

Teleoperation and Mobile Computing: Technologies used for mobile computing are discussed including XML, WAP and architectures suitable for teleoperation. Communications technology including the internet and the mobile phone network are considered. The course addresses supervisory control versus autonomous control and the tradeoffs involved. The significance of bandwidth is considered with an emphasis on approaches suitable for low bandwidth environments. Existing examples are presented, commercial opportunities examined, human interaction considered and social implications of the human-automation system discussed.

High Performance Scientific Computation: This course provides an introduction to High Performance Computing with an orientation towards applications in science and engineering. It addresses high performance computer architectures including vector and parallel processors and describes how an algorithm interacts with these architectures. It also addresses aspects of numerical computation in floating point arithmetic and fundamental numerical algorithms.

3D Web Technologies: This course introduces 3D technologies for the web, in particular, VRML and X3D. It examines the evolution of the web from its early beginnings as text only to the emerging rich media environment of text, 2D images, 3D graphics and audio. It looks at the concept of a scene graph and many fundamentals of 3D graphics, including geometry, transformations, lighting and shading. Finally, it considers applications of 3D web technologies.

6 Examples of projects

The following are brief descriptions of some eScience projects undertaken at ANU and RMIT. It can be seen that several of them have a similar orientation to those UK e-Science projects described above.

- Calculating and Visualising Transcription Factor Binding Sites in Genomic Data.
- Integrating River Condition Indices in environmental data.
- Social data submission website for the Institute of Health and Welfare.
- A simulation, and data archival tool for archaeological excavations.
- An interactive sound-scene spatialiser (a direct-manipulation interface for composing spatialised music).
- A 3D computer animation tool for a virtual reality theatre.
- A 3D drawing tool for a virtual reality theatre.
- Development of the Visual Python programming package and porting to a virtual reality theatre.
- Integration of a Data Mining package with a visual browser for scientific data.
- Development of a GUI interface for a tool for maintainence of multimedia, multilingual websites.
- Development of a cooperative, virtual environment application (together with a French research institute - the student worked in France!)
- The use of XML, including XHTML, MathML and X3D for scientific web publishing.
- Visualisation of isosurfaces using Java3D and the Marching Cubes algorithm.
- Sofware-based edge blending in tiled projection walls.

7 Graduates

At the time of writing, we are aware of the job placements of some 14 eScience graduates. Five of these have found employment in the university sector or CSIRO, 3 are employed by "high-tech" industry in positions which could be called eScience, 4 are employed in the public service and NGOs (1 new position and 3 changed positions as a result of eScience qualifications) and 2 have gone on to further study. About 8 students have withdrawn from the eScience programs and another 4 are recent (first-semester 2004) graduates looking for work.

These graduate destinations justify the following observations:

- Many jobs obtained appear to have an "eSciencey" character.
- Job opportunities have mostly been outside of industry.
- We have yet to evaluate the success, or otherwise, of international students competing for positions in Australia and overseas.

8 Disappointments

The eScience:Applications and Technology project has had many successes and some disappointments. The first disappointment has been the lack of industry interaction. Although the project management committee had 3 industry representatives, and although an eScience trainee-ship scheme was briefly entered into with one IT company, industry interaction has largely been a casualty of the internet bust. In retrospect, it is hard to imagine a more disastrous time to have begun an educational project aiming for industry involvement: financial pressures forced a discontinuance of the trainee-ship scheme and our one trainee was made redundant!

Our second disappointment was the ultimate lack of viability of the RMIT eScience program. In spite of extensive advertising in local newspapers, science teachers newsletters and at RMIT postgraduate information sessions, the program did not attract enough students to be self-sustaining. Contributing factors to this lack of success were

- many students with science and engineering backgrounds preferred to undertake other RMIT postgraduate coursework programs as they saw these as more directly providing opportunities to move into the IT industry
- the general IT downturn, and
- the cost of postgraduate coursework.

However, a small number of RMIT students were attracted to the true notion of eScience and its blend of programming, IT and computer skills along with courses which had a more scientific flavour. This was particularly true of people interested in Bioinformatics and Computational Biology. Unfortunately, ultimately there were not enough of these students.

9 Successes and challenges

The ongoing educational programs at ANU are the big successes of the the eScience initiative.

As remarked above, the eScience graduate coursework programs have operated in an environment of volatile student demand - not only in terms of student numbers but in the nature and motivations of the student body. Because we had the luxury of a government grant, we were able to tune the initial program offerings on a small number of students. These students were all domestic and were mostly from a science and engineering background who were mostly looking at converting their skills in the direction of employment in the IT industry.

After the introduction of the MIT, in 2002, we also attracted a number of domestic, part-time IT professionals as well as one or two international students. The motivation of the IT professionals was mainly to "upskill" in modern aspects of IT with a view to career advancement or just for personal interest.

From 2003-4 there has been a large influx of international students from the Indian subcontinent and China. A majority (perhaps 60%) of these students already had degrees in IT. Conincident with this change in student profile, academic results for eScience students have shown a distinct "double-humping". In courses where a reasonable comparison can be made, there is a group of eScience students (domestic, professional and international students amongst them) who compete effectively with our best undergraduates. Then there is a group of students who succeed at pass and credit level. The tail of this distribution also includes some students who fail subjects; sometimes repeatedly.

Our response to changing student profiles has been to tune the program offerings on a regular basis; in fact there have been revisions to the program structures and course offerings on a six-monthly basis. In the beginning, these tunings have been in the direction of liberalising the program structure. In recent times, we have begun to ephasise the "eScience" nature of the coursework offerings again. In particular, students who already have a strong background in conventional IT are directed towards our "eScience" options or towards other free electives in science, engineering or statistics. We have also introduced a new, elite, project structure which will allow talented students to effectively compete for PhD scholarships on the basis of their project work.

The growth in international student numbers has put pressures on domestic resourcing and on the abilities of lecturers to deal with special study skills and language issues. Our response to this has been to turn to the Academic Skills and Learning Centre on campus and to initiate some active collaboration with them (as well as directing students to them on an individual basis). For example, in 2003 and again in 2004, one of the core, compulsory MIT(eScience) courses has undertaken a collaborative exercise with ASLC to give feedback to students on their assignment work as well as on the examiner's marking. Indications are that this exercise was successful as well as cost effective.

From the beginning, we have attempted to be flexible in the times that ANU eScience courses were run and several have had lecture blocks in evenings and on weekends. This is very much appreciated by our part-time students but can be stressful on staff and on the full-time international students (who would appear to favour a regular, three-times a week, lecturing model). The best solution would seem to be to offer both models simultaneously and we will look into this for 2004.

As mentioned above, several courses initiated as part of the RMIT eScience program continue to be offered and attract significant student numbers. Facilities established at RMIT by the eScience grant, including a portable, tiled projection wall, continued to be used in a number of student projects.

10 Postscript

In 2004, following a high-level delegation to study the UK e-Science program, the Australian Research Council published a discussion paper (Australian Research Council 2004) which outlined a vision for, and flagged a new pilot support program in, "e-Research" (possibly used in preference to "e-Science" in order to be more inclusive of the humanities). The paper points out that the total, world-wide investment in e-Research activities (including grid technology) is over A$3 billion. It would seem that Australian universities will hear a lot more about e-Research in the years to come, and it could be that our experiences may provide some guidance for others wishing to set up similar education programs. More details about our programs can be seen from our web sites at (Australian National University 2004) and (RMIT University 2004).

References

Australian National University (2004), 'ANU eScience website'. http://eScience.anu.edu.au. Last accessed 8 June 2004.

Australian Research Council (2004), 'A Pilot in e-Research support'. http://http://www.arc.gov.au/publications/arc_publications.htm#2004. Last accessed 8 June 2004.

Research Councils of the UK (2004), 'RCUK e-Science web site'. http://http://www.rcuk.ac.uk/escience/. Last accessed 8 June 2004.

RMIT University (2004), 'RMIT eScience website'. http://www.cs.rmit.edu.au/courses/postgraduate/escience/. Last accessed 8 June 2004.