

# Scholarly Information Futures: A Faculty 10-year view

## Submission from the Faculty of Science

Access to and maintenance of various types of knowledge and information underpins many of the research and teaching activities in the Faculty of Science. The following details areas that the Faculty has a particular interest in maintaining, enhancing and/or developing in the future.

Martin Seviour (*Physics*), Tim Dyce (*Physics*), Janet Hergt (*Earth Sciences*), Raoul Mulder (*Zoology*), Jon Pearce (*Information Systems*), Andrew Robinson (*Mathematics and Statistics*) and Spencer Williams (*Chemistry*)

### ***1. Research:***

#### **a) Computational Resources**

The School of Physics, the School of Chemistry, The Department of Mathematics and Statistics and the Department of Earth Sciences are heavy users of computational resources and data storage for research purposes.

Within Physics, the Experimental Particle Physics (EPP) group is a pioneer of GRID computing in Australia and currently operates a 100 CPU node Tier-2 GRID facility with 150 Terabytes of storage. They plan to grow this facility over the next 5 years to provide 500 CPU's with close to 500 TeraBytes of online storage. The University of Melbourne houses this facility within its Queensberry Street data centre. While the provision of the physical infrastructure has been very helpful, the group is stretched to maintain this facility. The local Tier-2 turn gives access to the 100's of petabytes of data generated by the CERN experiment which is distributed throughout the world and allows the group to be full partners with their international colleagues.

The Physics Astro group is also a heavy user of computational resources. Their computational needs are expected to grow significantly as their participation in LIGO Gravitational Wave detector scales up and from when data processing from the Square Kilometer radio array commences. In both cases the Astronomy group foresees the need to become heavy users of GRID facilities. The scale of both these endeavors is similar to those of EPP.

The Faculty of Science also makes extensive use of conventional large scale computation facilities.

The Physics Centre for Excellence in Quantum Technology, housed in the MARC group, makes extensive use of large scale computational facilities for large scale solid-state device modelling and testing of quantum computing architectures and algorithms. Currently these needs are met via the use of dedicated large scale facilities and access to the world's largest supercomputers via international contacts. This Centre will be the nucleus of a University wide facility for materials related research. Storage and visualization requirements as a result of large scale ab-initio structure calculations (e.g. Density Functional Theory, million atom tight-binding nanoelectronics) are also high. Research groups in Chemistry employ Quantum Chemical calculations that can effectively utilize 100% of the resources of any facility provided. The Department of Mathematics and Statistics (Maths & Stats) research groups employ high performance computing for research in the fields of Applied Statistics, Statistical Mechanics &

Combinatorics, and Continuum Modelling. Finally researchers within the Department of Earth Sciences employ the large scale computer facilities for global climate modelling.

The Physics research groups mentioned encompass approximately 50% of the research strength of the School of Physics and attract in excess of 3 million dollars of annual research funding. For all 3 groups, high speed Internet access is vital as are large scale computational facilities and data storage. Both Astro and EPP utilize GRID to gain access to a world-wide facility hundreds of times larger than could be obtained locally.

The Department of Zoology reports that research data will need to be submitted to a National repository.

In addition to the above, the Faculty of Science has a general need for online storage and backup of research and teaching data. This is particularly acute in the Department of Earth Sciences, where Journals demand supplementary data before they'll publish some articles. This data must be placed online at a publicly accessible location.

Large volumes of critical research data currently reside on volatile mobile drives, the failure of which frequently causes large productivity losses and data loss.

Storage infrastructure to meet research needs has great potential for cost savings across the University, since at current the solutions purchased by individual research group lack scalability. With the bulk of the purchase outlay going toward control infrastructure and not disk.

A SAN provided at the University layer could capitalise on economies of scale, by combining the funds wasted in control infrastructure across the university into a single system. Leaving researchers free to purchase disk a known cost, with the added benefit of easing grant applications.

This kind of storage offering BYO server basis would allow researchers productivity, but not impinge on application functionality. Since researchers need only bring their own server and pay for disk to get started.

Such a facility would also researchers with data stored on legacy hardware to migrate to a modern facility and which can be archived for future reference.

### **Recommendations:**

- That the University fund a group of system administrators to maintain a collection of gLite based GRID facilities. This is the cutting edge of a new enabling technology that allows local researchers access to facilities hundreds of times larger than could be provided from just local resources.
- The University should fund a central facility (SAN) for scholarly data storage. Individual departments could pay for their own disks housed within the facility. Such a facility would significantly reduce the duplicated overhead as each department implements their own local solution to data storage and backup.

- The University should maintain or enhance its ability to access large scale computation resources either locally or remotely.
- Access to very high speed, 10 Gbit/sec or greater, Internet connectivity is vital for a number of high profile research activities in the Faculty of Science.

## **b) Access and production of knowledge:**

There is a broad view in the Faculty of Science that supports the continued expansion of access to online journals and books. A particular area of interest is ensuring access to complete archives of online journals and online versions of books. Increasingly, many 'monograph' type resources are available as online subscriptions, and are continuously updated, whereas the print versions are only released on an infrequent basis.

The University's current access to online journals, through a complicated series of web-maneuvers specific to individual papers or to specific journals, does not match world's best practice. At many institutions around the world, every computer attached to the University network has full access to all online journals. So a researcher need only search in "Google Scholar", obtain the link to the Journal and directly access the article in question.

The School of Chemistry has specific requirements for online access to both Journals and databases.

The Department of Zoology is keen ensure that they are consulted when subscriptions to specific Journals are contemplated as are Maths & Stats and Physics.

Over the past few years, a movement towards direct digital publishing and Open Access has gained strength. This promises to significantly reduce the cost of Journal production. Much of the strength for the Open Access initiative comes from the use of the <http://arxiv.org/> preprint server. By publishing pre-prints on arxiv.org scholars can clearly establish priority as well as enable open access to ideas and results. Physics and Maths & Stats researchers have been at the forefront of this activity. In part this has been facilitated through the widespread use of the Latex markup language for scholarly publications in these disciplines and on arxiv.org. This is a freely available open source standard and has significantly reduced the cost of online publications. Over 90% of Physics and Maths & Stats graduate students are proficient in the use of Latex and we provide introductory courses to help our students get started. There are now a variety of open source tools that simplify the production of high quality Latex documents which would enable scholars in other fields to have access to this powerful enabling technology. As more disciplines employ arxiv.org, the Open Access initiative will gain strength and provide a means to reign in the cost of Journal subscriptions.

We will always require access to printed books, particularly text books for courses, however their importance has significantly diminished over the past 10 years.

There remains a strong view of some within Maths & Stats that conventional printed Journal collections be maintained on site. Maths & Stats reports that it is well served by its current branch library which it sees as a substantial local asset.

The University could promote both its own knowledge base and encourage the widespread growth of knowledge through an online repository of Ph.D. theses. While some restrictions may

be placed on commercial-in-confidence sponsored research, the vast majority of Ph.D. theses should be placed online for all the read.

**Recommendations:**

- That the university retain, simplify and expand their access to digital journals. The University should also look to provide direct access to all online journals within the University network.
- The Departments within the Faculty of Science need to be consulted when decisions to purchase specific subscriptions are made to ensure that their needs are met.
- While printed Journals are important for archival purposes, these are of lesser importance than online Journals now.
- The University should encourage the growth of direct digital publishing. To this end it should train staff in the use of the Latex markup language and provide tools which facilitate its production.
- The University of should establish a digital repository of Ph.D. theses and make these universally available, except where this conflicts with commercially sponsored study.

**c) Subscriptions to software**

The Faculty of Science makes extensive use of some commercial software packages. To date the University has been rather poor at negotiating site-wide licenses for such products.

**Recommendation:**

- We support the University investigating whether significant costs could be realized through an institutional based subscription.

**d) Connectivity**

Many members of the School of Physics make routine use of video conferences to talk to collaborators in Australia and throughout the world. We expect that this trend will continue and requires access to high speed Internet connection. Other departments in the Faculty are not so well served and have found it difficult to acquire the funding to setup the facilities. In particular the Department of Earth Sciences has not been able to setup department-wide wireless access. Similar frustrations were identified in Mathematics and Statistics.

**Recommendation:**

- We support the continued investment in high speed Internet connectivity as being a core asset of the University in the digital age.
- This support should be provided all the way to the desktop of Scholars and to teaching spaces.

## ***2. Teaching***

### **a) Online delivery of course materials**

Essentially every undergraduate course delivered by the Faculty of Science relies on online delivery of teaching materials, and in some cases makes use of content rich multimedia capabilities. While the Learning management system adopted by the University has brought consistency to delivery of learning resources, it is clunky and does not allow the integration of even slightly sophisticated capabilities (eg Flash, various Freeware plugins etc).

The Department of Information systems foresees the use of emerging information technology to enable new levels of collaboration and connectivity. In particular it is now inexpensive to provide every student with their own online presence within the University of Melbourne network site. Doing so would help propagate the value of the University of Melbourne brand and encourage post-graduate loyalty.

The development of GRID technologies together with other large scale data repositories opens the possibility of employing real data in various teaching scenarios.

#### **Recommendation:**

- We support the continued expansion of the capabilities of the Learning Management System so as to support more interactive functionality.
- Ask the University to encourage the exploration of emerging information technology for innovative teaching practices.

### **b) Teaching spaces**

While uniform access to data projectors to allow computer based multimedia content is now standard in University Lecture spaces, in some lecture theaters this has made it impossible to use whiteboards or blackboards. This significantly decreases the quality of a lecture experience as the development of equations and the explanation of Science principles is best done live on a whiteboard or blackboard through the development of equations and the use of diagrams sketched by the lecturer. This is born-out through quality of teaching surveys.

The Department of Mathematics and Statistics reports that audio/visual facilities in some lecture theaters have deteriorated leaving the lecturers of new University breadth courses in a very difficult position.

The Departments of Maths & Stats and Earth Sciences would like to setup fully networked and videoconferencing teaching spaces but have been unable to do so through lack of funding. The department of Information Systems would like to investigate the use of emerging collaborative technologies to enable remote access and to break down the distinction between lecture, tutorial and workshop spaces.

**Recommendation:**

- The Faculty of Science strongly urges that the University retain full access to whiteboards in Lecture theaters and to maintain the audio/visual facilities within them.
- The University should provide a means to enable the use of videoconferencing and a fully networked environment to construct innovative teaching spaces.

### ***3. Infrastructure.***

The University is currently undertaking a substantial upgrade of its network infrastructure. However current plans have this excellent new service terminate at the door of building but not provide what is needed for the internal networks within the department.

There is not much good having superb policies on high speed wide bandwidth communications, but the last link to the users remains chaotically managed and of uncertain adequacy!

There are a variety of opportunities for innovative new personal hardware designs to improve the productivity of scholarly activities. In particular, wireless tablet technologies have finally become mature enough to be useful. These may provide an alternative to paper based input, particularly if the ability to translate mathematical equations to Latex or some form of computer code can be perfected.

Currently the University only supports the use of Windows and Macintosh client operating systems. However Physics, Zoology as well as Maths & Stats make substantial use of Linux clients as well as Linux servers. Physics finds that the the support requirements of Linux desktops are approximately one quarter that of Windows machines so the one staff member can support 4 times as many Linux desktops as Windows desktops. With the worldwide move to non-proprietary and open data formats, the University could obtain substantial cost reductions in both License fees and support costs by utilizing Linux as a fully supported client platform across the University.

All departments in the Faculty of Science value their local IT support teams but this is particularly true of Maths & Stats, Chemistry and Physics. Here the IT teams have acquired critical mass and have established local experience with the quirks and requirements of each department. They are now a vital part of the "corporate culture" and support base of their disciplines. While cost savings can be made through shared facilities, the Faculty of Science needs to retain local departmental IT support and experience. In addition local support enables a diversity of approaches to Scholarly knowledge within the University. At some level such a diversity provides an insurance policy as the evolution of digital information remains vibrant, rapid and uncertain.

**Recommendation:**

- The University should facilitate the creation of excellent internal department networks.
- Academics are encouraged to investigate new hardware devices as a means of improving productivity.

- The University should support the use of Linux on clients as appropriate.
- Departmental IT support will be required to support the specific needs and requirements of departments. This will also enable some level of diversity of approaches to scholarly knowledge within the University.