

# Changing user behaviour using a digital repository system

Martin Borchert  
Associate Director, Access Services  
Information Services  
Griffith University  
[m.borchert@griffith.edu.au](mailto:m.borchert@griffith.edu.au)

## ***Abstract***

*Griffith University is developing a digital repository system using HarvestRoad Hive software to better meet the needs of academics and students using institutional learning and teaching, course readings, and institutional intellectual capital systems. Issues with current operations and systems are discussed in terms of user behaviour. New repository systems are being designed in such a way that they address current service and user behaviour issues by closely aligning systems with user needs. By developing attractive online services, Griffith is working to change current user behaviour to achieve strategic priorities in the sharing and reuse of learning objects, improved selection and use of digitised course readings, the development of ePrint and eScience services, and the management of a research portfolio service.*

# Introduction

Griffith University has been developing its digital repository service using the HarvestRoad Hive repository system, and aims to provide a range of integrated repository services that span the learning and teaching, library, research and administrative functions of the University. This paper describes a number of digitisation projects underway at Griffith University, including learning objects, digitised course readings and university intellectual capital (ePrints, Research Profile, and eScience), and how the digital repository system is being developed to force new workflows and encourage user behaviours to achieve institutional goals in learning and teaching, and in research. Current operations and issues related to digital collections are discussed, with solutions presented by the application of the digital repository.

Griffith's digital repository presents many opportunities for improved workflows, storage, management, access, reporting, and re-use of the University's digital collections. The development of a repository system is expected to empower academic staff in these processes and offer them convenient online services prompting changes in demand for service and user behaviour.

## Literature Review

Much has been written on the use of digital libraries and their usability, however the use of digital repository systems, whether for learning and teaching, library or other applications, is a newer field with little research available to date. Prior research into the usability of digital library services, and the motivation of users is summarised. General comments and principles about usability can be applied to the use of digital repositories. The concepts of ease of use, and alignment to user needs are cornerstones to the development of Griffith University's digital repository project.

The well known IT saying 'If you build it will they come?' (Markus and Keil, 1994) remains highly relevant more than a decade later. The success of IT projects depends not only upon the IT investments, but also upon the alignment with business needs and processes. Even the best aligned systems are not used if their creators do not consider user's motivations and commitments (Markus and Keil, 1994).

User motivation can be discussed in the context of commitment, compliance, identification, and internalisation (Malhotra and Galletta, 2004). Users will avoid learning new technologies when they do not perceive direct personal benefits. It is essential to balance technical usability (from an IT designer's perspective) with performance usability (from the users' perspective). Organisations should harness employees' sense of pro-social attitudes, tendency towards good corporate citizenship and desire for peer recognition, especially in regard to optional use of new systems. The ideal situation is where users internalise a desire to use a system to empower themselves to achieve their own objectives. Griffith is attempting to harness the corporate goodwill of academics by building the repository system as a fully integrated supporting technology for the already well established signature service called Learning@Griffith. The repository is thus presented as a tool to empower academics to more efficiently achieve their goals in delivering online learning experiences.

Behaviour modelling and applied scenario based testing of new systems can be compared and it is suggested a process that combines both is required to globally test a whole system by testing the use of each of its components (Uchitel, Kramer and Magee, 2004). Prior user testing of the Learning@Griffith system will be supplemented by user testing of new digital repository components.

Interest in the use and usability of learning object repositories has grown and has attracted funding within the Joint Information Systems Committee (JISC, 2005), UK, for a study into the Community Dimensions of Learning Object Repositories. This study aims to identify and analyse the factors that influence practical uptake and implementation of learning object (LO) repositories, with a focus on social and cultural issues in support of programme objectives. The way users interact with learning objects and how they search for them on the ARIADNE repository, and an analysis of system log files to track usage, have been studied, with findings that searchers predominantly used default search settings and used less than half the data elements available (Naijar, Ternier, and Duval, 2004). Models for predicting and delivering timely and appropriate access to documents within repositories based upon past access log files have been developed (Recker and Pitkow, 1996). Griffith wants to keep search interfaces simple yet capable of being searched via Griffith's federated search service called CrossSearch. Griffith will also be in a position to analyse search log files once the learning object repository is functional and in use.

The investigation of the barriers to change in developing open access repositories such as institutional ePrint services has been discussed in the context of the legal framework, information technology infrastructure, business models, indexing services and standards, the academic reward system, marketing, and critical mass. Issues pertaining to the academic reward system within universities, and related to the relative prestige of publisher managed and open access services, were identified as barriers in comparison to legal and technological issues. It is suggested academics be paid a reward to deposit their papers into their institutional repository, to add greater incentive for researchers to participate (Bjork, 2004). Griffith is trying to learn from the experiences of other institutions regarding the promotion of the benefits of ePrint services to researchers before making such a system locally available. By promoting the benefits of open access to researchers' academic profiles, Griffith will attempt to short-cut the long content contribution curves experienced by some institutions

## **Digital Collection Projects at Griffith University**

Griffith University is developing a number of digital collection projects using the repository system. Although development of the underlying repository system is well advanced, the first repository services are yet to be made available to clients at the time of writing (November, 2005) This paper describes the current operations, issues to be addressed and possible applications of the digital repository system to learning object collections, digitised course readings and past examinations which are being introduced for early 2006 and also institutional intellectual capital collections including ePrints, research portfolio project, and eScience data archives which will be introduced as a second stage during 2006 and beyond.

## **Learning Objects**

Higher education sector requirements for flexible and online learning environments and student centred learning have required Universities to implement online learning management systems (LMS) and populate them with a wide variety of learning objects, developed locally by academics and educational designers. Learning objects are defined here as digital objects providing a self-contained learning experience or fulfilling a discrete learning objective and are comprised of a number of sub-components called digital assets. Universities face the problem of developing systems that promote greater coordination, sharing and re-use of learning objects by academic groups and across the curriculum in order to extract greater value for money from their investments in the development of learning objects; the goal being to transform disorganised and hidden aggregations of learning objects into organised, accessible and sharable and reusable collections.

### **Current operations and issues to be addressed**

Griffith University has invested heavily in the delivery of an integrated online learning environment called Learning@Griffith which uses the Blackboard LMS. Griffith has also made the creation of learning objects a priority in order to enrich courses and provide students with quality interactive learning experiences. Four teams within Flexible Learning and Access Services, combining educational designers, faculty librarians, graphic artists and web developers have been created to provide client-focus, subject specialist professional services to assist academics in the creation and delivery of online course environments and materials. These teams work closely with academics to develop resources and to assist academics to develop resources themselves. Much work is done on a project basis with funding available for targeted education specifications and required outcomes.

Resources developed by the teams and by academics alike are uploaded to Learning@Griffith and made available to students via online courses. All academics and students have access to all online courses and attached learning objects. Online courses for the prior semester are recycled for the next semester, and thus academics may have two versions of any online course available to them.

The issues of access, sharing and reuse and archiving are complex primary drivers for the establishment of a digital repository service supporting learning object collections. Some of the aspects of these issues are listed here.

Repositories may address some of the weaknesses of LMSs without forcing a new system upon users. Some of the specific issues include:

- LMSs do not provide user friendly search systems across learning object collections, while repository systems can provide such search functionality;
- An institutional LMS can develop into a large system, making the development of functional virtual communities difficult;
- Repositories can support the design and application of suitable metadata schemas supporting the description of, and access to, learning objects across the University;
- Digital repository systems can be implemented as a hidden system behind the LMS interface, thus not requiring academics and students to learn a new interface.

A major benefit of repositories is that they group learning objects from across the LMS into specific collections designed to aid sharing and reuse. Issues include:

- While academics usually want to share their research, they often regard their teaching portfolio as personal, and so do not share the learning objects they have created with other academics;
- There exists no environment for peer review of learning objects created by academics;
- Similarly there is no environment for academics to actively share existing learning objects;
- Learning objects created by educational designers are delivered direct to the academic's online courses, and are not necessarily promoted to other academics who might use them;
- Academics will duplicate effort in developing learning objects rather than reuse existing similar objects.
- Each object may be linked to from multiple courses thus negating duplication of objects within the LMS.

The LMS may provide versioning of courses across the academic calendar but does not support versioning and archiving at the individual learning object level. Issues include:

- Using the LMS as a delivery system, academics may delete unneeded objects rather than store them for later;
- When updating objects, academics may delete or lose the original version and so version control is lost;
- Academic staff turnover results in the loss of valuable learning material because outgoing academics may delete material or take it with them. Alternatively, incoming academics may delete existing materials when they apply their own materials and style to a course;
- Academics who delete material can ask for copies to be provided from archive tapes. This is a highly manual and costly process which does not empower academics to help themselves;
- Repository systems can allow for long term archiving of courses for legal purposes.

### **Applying the digital repository to learning objects**

One of the first applications of Griffith's new digital repository is the management, archiving and access to learning objects. Hive is being developed with bureaus (structured collection spaces) to perform these functions separately and 'hidden behind' the Learning@Griffith interface. To make the repository as seamless as possible Griffith is developing a number of application programming interfaces (APIs) or building blocks to link the Learning@Griffith interface with the functionality provided by the digital repository. Academics have been using various building blocks for years, and it is expected academics will take to newly developed building blocks designed to support sharing. Building blocks are pieces of software designed to provide specific functionality within the LMS, which can also be used to connect the LMS to the repository.

Mirroring the academic and flexible learning support team structures, a category is being developed for Arts, Education and Law; Business; Health, and Science, allowing academics and educational designers alike to deposit objects within subject specific collections to facilitate access, sharing, peer review, and re-use in subject based virtual communities. The development of a suitably customised metadata schema is critical to the success of this aspect of the project. A new schema has been developed in consultation with academics and the

Flexible Learning support teams for the purposes of description, version control, rights management, access and preservation.

Initially, objects developed by the Flexible Learning support teams, identified as being of high economic value (based on cost of development and re-usability) are being divided into subject collections and loaded to the repository. Description of this retrospective collection of about 500 high economic value objects (in the first instance) requires expert understanding of the content and educational context of each object. Educational designers will be required to develop the metadata for objects developed by their teams, and to deposit the metadata and the associated objects into the repository to develop a best-practice example for academics (and the support teams) to use and benchmark. Academics will be able to access this collection, and link to objects from their online courses, but will not have rights to delete objects in this collection. These high economic value subject collections will be centrally managed and highly controlled.

Another category is being developed for digital assets. These are highly re-usable, create efficiencies in the development of learning objects, and are well suited to the digital repository. This collection will be highly used by the support teams, but will also be available to academics.

Objects created by academics could be placed into the repository via four alternative workflows under development. The first is by academics requesting that objects be placed directly into the appropriate high economic value subject collections, and linking to these from their online courses. The second is by academics placing objects into 'uncontrolled' subject collections, and then linking to objects from their online courses. The third is by Digital Repository and Flexible Learning support team staff identifying high economic value objects created by academics, and negotiating with academics for these objects to be copied to the centrally managed and highly controlled collections. Unfortunately there is no means within Learning@Griffith to force academics to place learning objects into the repository rather than load objects directly to their online courses within Blackboard (as is the current practice), and this creates the need for a possible fourth workflow by which support staff will identify valuable objects to again be copied into the appropriate subject collection within the repository.

For the purposes of providing longer-term archiving of courses, as well as access for academics to past versions of courses (and objects), Griffith will zip courses for each semester and deposit these packages into the repository. It is expected a four year rolling archive will be kept, as this time period provides maximum returns, with most students completing most academic programs within four years.

## Digitised Course Readings

Griffith currently provides access to digitised course readings via its GriffLink library catalogue. Major problems with the system are that all readings are available to all students in all courses all the time. This creates problems because copyright compliance issues may prevent new readings being made available from books and journal issues for which a reading is already available online. Version control is also not available. By creating an integrated environment in which academics can best select and make readings available to be appropriately 'pushed' to relevant students within the Learning@Griffith environment, it is expected Griffith will make optimum use of existing staff resources, existing digitised

readings, and library databases and ejournal collections. Griffith's current systems and workflows for the provision of digitised course readings are discussed in some detail here. Inherent problems are expanded upon, and opportunities for solutions offered by the digital repository implementation are developed.

### **Current operations and issues to be addressed**

The transformation to online learning environments has necessitated university libraries digitising their existing hardcopy collections of selected journal articles and book extracts selected by academics and commonly referred to as course readings. Over the past seven years, Griffith has digitised 12,000 course readings available via the Griffink library catalogue, with links integrated into Learning@Griffith which are automatically generated using a search script based on the course code.

The Digitisation and Distribution Team is a centre of expertise in digitisation functions, and is a single source of contact for academic staff requesting the service. Copyright compliance of all learning and teaching materials is the responsibility of academic staff at Griffith University. One of the Digitisation Team's major tasks is to assist academics with copyright compliance so that only compliant materials are used, and made available to students under the Copyright Agency Limited (CAL) Licence. Digitisation staff advise academics when selected resources are non-compliant before they are digitised, so alternative sources can be selected.

Each semester, the Digitisation Team receives approximately 300 requests from academics for new readings to be digitised. These requests are typically made by academics with minimal lead time to the start of semester and continue to a lesser degree throughout the whole semester. A service level agreement between the Digitisation Team and academic clients requires requests for digitised readings to be made at least four weeks before the start of semester, and that requests for readings will be processed by the Digitisation Team within five working days. The service also requires that academic staff provide a print copy of the required reading with their request. Digitisation and Distribution staff also assists academics in the production of print course packs consisting of lecture notes and also some course readings. By combining these functions into one team, the Digitisation staff are able to actively flag course readings produced in print format, to be replicated or diverted to digital format (copyright compliance permitting).

The following issues surrounding digitised readings need addressing at Griffith:

- With limited staff available, demand for digitisation services during peak periods often exceeded the team's ability to provide services. There are not the resources available for additional staff for digitisation work, suggesting Griffith University needs to develop an alternative, more efficient procedure;
- Supporting academics with copyright compliance is a time consuming and complex task. Academics may send requests for readings to be digitised that do not meet copyright compliance requirements. Digitisation staff check existing holdings and negotiate alternative arrangements with academics. It is a difficult task to raise academics' awareness of these issues, limitations and procedures. An automated system that replaces the need for academics to apply knowledge to the selection of readings would provide a more efficient and timely service;

- Not all academic staff are using the service, and thus benefiting from the copyright compliance support services offered by the digitisation and distribution team. If all academics were to use the service, then greater efficiencies would need to be found using workflows generated by the digital repository in order to cope with the additional demand;
- By placing readings in the library catalogue, these readings have been provided on a permanent basis because readings are not 'published' and then 'unpublished' as the need for particular readings changed. This has made it impossible for Griffith to provide all required readings across all the University's courses within the requirements for copyright compliance (within one chapter or 10% of a book, or a single article within an issue of a journal, unless they are on the same subject matter);
- By leading students from the Learning@Griffith online interface to the GriffLink interface to access readings students are working within two (albeit integrated) environments;
- Academics often choose older materials to be digitised. There are of course valid pedagogical reasons for doing so in some cases, but evidence suggests that academics are choosing new print readings to be digitised without first considering existing alternative reading sources available from databases and ejournals. Only a small percentage of requests made for readings have been linkable to library databases;
- Many academics are linking to online articles directly from Learning@Griffith course sites, thus creating a second point of access to many course readings. This has the possibility of creating a non-uniform and possibly confusing environment for students. It also makes it difficult to determine the number and nature of links to the range of library databases, this being important information should Information Services consider cancelling subscriptions to particular databases;
- Library databases and ejournals (costing over \$3 million dollars per annum in subscription fees) have not been leveraged to their full potential for the purpose of learning and teaching;
- The system of digitised readings relies heavily on the Copyright Agency Limited Academic Licence as the basis under which readings are digitised and made available. The greater the number of readings digitised under the licence, the more reliant Griffith University is upon the licence.
- The library catalogue was effectively used to create reports for a CAL Electronic Use Survey (EUS) in 2004, but generating reports based on data drawn from MARC records and then manipulating this data into the EUS reporting template was a tedious task. A better system is required.

### **Applying the digital repository to digitised course readings**

By developing the repository to manage digitised course readings, Griffith University is developing solutions to the problems outlined above. Workflows within the repository will allow academics to request new readings, search for existing readings, and book copyright compliant readings for specific times. The repository will provide for copyright violation detection in the booking system and will provide workflows for the creation of new readings.

A secure bureau within Hive has been developed specifically for the storage and management of course readings. Readings digitised (created) in each year will be stored within a specific category for that year. Within each category, readings will again be subdivided into subcategories based on publication format such as book extract, journal article, conference paper, reports, grey literature (pamphlets etc). These divisions will assist staff to manage the

collections, and report on their creation and usage for CAL Electronic Use Survey (EUS) survey purposes.

Griffith's collection of 12,000 course readings is being bulk migrated to the digital repository. Data from existing MARC records in the library catalogue is being extracted in MARC XML and parsed against the Griffith Metadata Application Profile (GU-MAP) in Hive XML format, with these records also being loaded to the repository. Client access to course readings will then be wholly moved from the library catalogue to Learning@Griffith.

Two application programming interface (APIs) have been developed. The first API has been created to assist in the selection, copyright compliance and management of digitised readings and has been designed with interfaces for both academics and Digitisation Team staff. To request a digitised reading, academics enter the system via their Learning@Griffith Course Administrator Account where an online form is presented requesting the bibliographic details of the required reading. Data is requested in standard format minimising the need for and possibility of, free text interpretation. The academic's request will be forced via a relatively seamless three step online workflow which has been so far partially designed to maximise usage of Griffith's library databases, existing course readings, and ensure copyright compliance.

Step 1 of the process is being designed as a mandatory step to map the bibliographic details of the academic's request to Griffith's SFX Open URL Resolver service where the resource is searched across Griffith's digital library resources. This is proving to be a technically difficult development. Where the book extract or more commonly the journal article, is found online, the URL and source is presented to the academic who may select the link to be copied (manually at this stage) into the resources list for their course. These URLs are ideal for this purpose because they are stable as long as Griffith continues to subscribe to the resource. Academics are not required to construct URLs according to an algorithm, or protocol such as Digital Object Identifiers (DOIs) or proprietary systems as used by a number of publishers. Where the book extract or more commonly the journal article, is not found online, the process progresses to Step 2.

Step 2 flows the details of the request to be searched across Griffith's existing digitised course readings. Any exact matches or near matches (such as a match on ISBN, ISSN, or combination of author and title) are presented to the academic in the search results. Where an exact match is found, the metadata, prior 'bookings' and a link to the reading is provided to the academic for perusal. More than one course may use the exact same reading at the same time. The academic checks the availability and prior bookings for the exact match reading, and near match readings (same edition of the book, or same issue of the journal) and confirms this is the exact reading required and is then presented with an online form requesting the dates the reading is required. It is important that information about the prior existence of near match readings is presented to the academic at this time because a prior booking for another reading from the same source may result in the requested reading being digitised but not made available (in accordance with copyright restrictions) to students during the required period. (Griffith is thus interpreting the copyright law and the CAL agreement such that it is the communication of a single digitised reading from a book edition or journal issue at one time and not the storage of a single digitised reading from a book edition or journal chapter that is the primary restriction.)

Where near match(s) only are found the academic is presented with the bibliographic details of, and a link to, these resources, thus allowing the academic an opportunity to determine whether a new reading is required or not. At this point the academic may select a near match reading and proceeds to Step 4.

Step 3 is used where no match is found. The academic is provided with the opportunity to place a request for a new digitised reading to be created. Instructions about copy source, copy quality, page layout and pagination are provided by the Digitisation Team. Bookings for readings can be made either within the current semester or one semester ahead following a flagfall date (example 6 weeks prior to semester to be confirmed) only. Bookings can be made for as little as one week blocks (with a separation period), but for no more than one semester period at a time. By entering the course booking system via the Course Administrator functions of Learning@Griffith, the system knows the course code for which the reading is booked.

A rights management system (incorporating copyright and licensing) is under development which will assist academics with copyright compliance of requested readings. Information about the reading and the required dates are checked against the database and the booking system. Prior bookings for chapters in the same edition of the same book, or articles from the issue of the same journal are flagged and presented to the academic. If no prior bookings are found the academic is asked to confirm the booking. If a prior booking from that source is detected for that specified time period, then the course code and the name of the academic for that course may be provided. The requesting academic thereby is provided the information required to either change the specified time period or contact the other academic to negotiate changes to the prior booking period to allow the newer booking to be made.

The second API has been created to work as a Blackboard building block linking the Hive bureau for digitised readings to the course pages within Learning@Griffith. Each online course contains a section presenting students with metadata and links to resources, such as lecture notes, past exams (also housed in Hive), web sites and digitised course readings. Within the resources section of each online course, metadata about each reading (including the dates it is to be available, as specified in the booking) is presented. This second API could interrogate the bookings data to provide active context sensitive links to appropriate readings during specified time periods only.

An issue under consideration at Griffith is whether to provide all readings for a course from one single location within the online course, or whether to present students with a list of course readings specifically for each week of the course. Pedagogical requirements present a strong case for readings lists customised for each week of the course. A flexible system is needed to serve the needs of academics and students alike. The bookings system of course means that links to readings could appear and disappear from an online course over time such that metadata about each reading is provided to students during the entire semester period, but the actual links to the readings are provided during the booked time period only for each reading.

Some time prior (about six weeks) to the start of the next semester, academics (Learning@Griffith course administrator) will be provided with the opportunity to book readings for the coming semester. The template will use the readings and bookings from the previous semester the course was available in as the start pointing, but no links to readings will be active. The academic will be required to reselect the use of each reading, and to

rebook the reading in time. Following the last day of each semester the booking system for that semester will be disabled and all links to all readings will be deactivated. The combination of disabling last semesters' readings and allowing pre-booking of readings provides an opportunity for academics to prepare readings well in advance of actual use. The system also encourages a competitive environment for booking readings, which will encourage academics to book readings (and make new requests) earlier, thus helping the Digitisation Team meet its SLA.

By storing and managing Griffith's digitised course readings in the digital repository, and by facilitating workflows and management functions using numerous Blackboard building blocks (APIs), Griffith is trying to provide mechanisms to improve the management of its course reading collection, and to change (primarily academic) user behaviour. The new processes will address the current problems associated with cost of service, return on investment, and compliance. The system will empower academics to help themselves; encourage better use of investments in library databases and ejournals; encourage academics to see and understand copyright law and the CAL agreement in operation and provide systems to adhere to restrictions in the use of materials; transfer the performance of certain tasks from digitisation staff to academics; increase the capacity of the system during peak periods; allow for a more flexible usage of digitised readings across time through a booking system; while retaining users within a single online environment.

## **Institutional Intellectual Capital**

Many universities and other research institutions have over the past five to ten years come to realise the benefits of centrally managing, providing access to, and promoting their intellectual capital. ePrints, the Department of Education, Science and Training (DEST) Higher Education Research Data Collection (HERDC) process, Research ePortfolio and eScience archive services are discussed, in the context of Griffith University's digital repository project. Issues and possible solutions are discussed with special reference to user behaviour.

### **ePrints**

Most implementations of institutional intellectual capital projects have so far concentrated on making the institution's research output (pre-print and post-print versions of published articles, and other materials) available via an open-access ePrint repository service. The benefits of such a service are well documented. A wide range of specialist (eprints.org) or general repository (Fedora, DSpace, Learning Edge, ExLibris Digttool) software exists from a variety of vendors providing repository systems, learning management systems and library systems.

Many universities in Australia and around the world have taken a lead over Griffith University in the development and implementation of an ePrint service. While some Australian institutions have thousands of papers in their repository (University of Queensland has 2064; Queensland University of Technology has 1500 at the time of writing this paper), many others have relatively few. Institutions with more successful repositories such as UQ and QUT for example, have seen an increase in the rate of deposit each year their repositories have been available, but this has arguably come about as a result of extensive promotion and at a significant staffing cost. Institutions with existing ePrint repositories have reported

reluctance on behalf of researchers and academics to deposit their research papers into the repository. In response, institutions have employed additional professional staff (usually librarians) to promote the service, obtain research papers from authors, and load these to the ePrint repository on behalf of authors. This can be a costly exercise, despite most ePrint software being specifically designed with self-deposit features in mind. Different institutions also have different policies on the purpose of ePrint repositories and the types of material collected.

## **DEST – HERDC**

Universities in Australia are required to report their research output to the Department of Education, Science and Training (DEST) on an annual basis via the Higher Education Research Data Collection (HERDC) process which is used to determine allocations to universities under various funding schemes. Griffith has to date collected and reported some 12,000 publications to DEST since 1996. In the past the process was completely manual, but a semi-automated process has been used for some years by which researchers deposit metadata about their research papers via the Griffith Portal (PeopleSoft), and also deposit the hardcopy of the post-print version of their work to the Research Publications Team. Griffith has had great success with this project, with a 24% increase in deposits in 2004 (up from 2003) and a 17% increase in 2005.

Issues related to the use of ePrints and the HERDC :

Awareness and commitment:

- Marketing and promotion of ePrint services within the institution to educate academics and researchers of the benefits, and requirements regarding publisher agreements;
- Researcher claims of lack of time to do deposit;
- Researcher ambivalence or lack of awareness of the benefits of ePrints;
- Researcher unwillingness to deposit their work into an ePrint repository in addition to other deposits such as to publishers and the DEST HERDC process;
- Limited staffing resources to employ additional staff to upload works to the ePrint service.

Fear and misunderstanding:

- Researcher fear of not being able to publish their work elsewhere such as publishers' journals (although about 90% of publishers allow this now);

Integration:

- Works suitable for deposit into the HERDC system are possibly a subset of works suitable (depending on policy) for deposit to the ePrint service;
- ePrint deposit requirement (generally) for the pre-print version of the article, but DEST HERDC deposit process requirement for the post-print version of the article;
- ePrints are possibly best gathered on a continuous process whereas HERDC research publications are reported on an annual basis, thus information required for reporting tends to be gathered on an annual basis also. This creates issues of timing and data flow between the two purposes;
- Many ePrint software choices are designed for self-archiving, but provide no bulk-upload function making retrospective loading of collections time consuming.

## **Applying the digital repository to ePrints and the HERDC process**

Griffith University is working to apply its digital repository system across the spectrum of institutional intellectual capital functions and systems. Although efforts have concentrated on the application of the digital repository to learning objects, digitised course readings and past exams in the first instance, progress is being made on conceptualising and planning the application of Hive to institutional intellectual capital projects.

A major workload and expense factor is how to affect author behaviour to encourage deposit of research papers. Griffith wants to leverage its HERDC success towards a newly developing ePrint service and is developing online services and processes that relate the two functions.

HERDC data collection via the Griffith Portal (an enterprise wide system based using Peoplesoft) presents itself as a natural extension to also make a deposit to the ePrint archive. By offering the two processes at the same time, it is envisaged this will create a more efficient and less burdensome task for authors. An issue that needs to be addressed is that the HERDC process requires the post-print version of the article, while the ePrints process generally requires the pre-print version of the article. To overcome confusion, the two requirements will be presented as linked, but separate workflows available concurrently from the same 'Research' screen of the Griffith Portal. The system needs to allow for alternative workflows to suit different types of works, and different HERDC and ePrint process timelines. It needs to allow authors to deposit a DEST recognised work to the HERD and ePrint systems at the same time, but it also needs to allow an author to deposit such a work (or various versions of a work) to the ePrint service at any time. The system must encourage the deposit of miscellaneous works to the ePrint which are not otherwise deposited to the HERDC process.

## **Research Portfolio**

Another issue in managing intellectual capital within large research institutions is providing a mechanism to store, manage and deliver information about experts and expertise within the organisation. For the purpose of this paper this will be called 'Research Portfolio' management.

## **Current operations and issues to be addressed**

Griffith currently manages and provides such information in a distributed way with each School, Faculty, Research Centre and the Office for Research maintaining this information using the phone book, and organisational web sites.

Issues include:

- Large institutions have difficulty managing such information and directing external clients to appropriate centres of expertise;
- Large institutions themselves and staff within them, have difficulty in identifying experts and expertise within the organisation;
- A single source of such directory information is required across the institution to satisfactorily direct external enquiries in a cross-disciplinary environment;
- Such a system should link into the institutional ePrint repository;

- Schools, Faculties, Research Centres will want customised views of the data (and sets of links to ePrints) for marketing purposes.

## **Applying the digital repository to research portfolio services**

Although research portfolio services are being developed primarily using the Griffith Portal PeopleSoft system, the digital repository provides a convenient means by which to provide repository services behind it. Linking between the research portfolio system and the ePrints and eScience datasets will be invisible and crucial.

The availability of the research portfolio is expected to create significant demand for the ePrint and eScience archive services because in effect, it is the marketing or branding layer that sits over the data archives. The portfolio service will allow individual researchers, Schools, Research Centres and Faculties to showcase their research, researchers and expertise in a highly customised way. This multi-layered and client-focused approach to marketing the University's research efforts is expected to generate high levels of client uptake and acceptance for this service, and to promote the ePrint and eScience services. Researchers want to present their research in a customised way that adds the most value to their work, and provides the most meaning to target audience groups. They need to wrap the products of their research in value-adding information about themselves, their peers, their projects and areas of expertise, as well as information about their respective Schools and Centres. Rather than presenting Griffith's research as a homogeneous whole (which of course it is not), encompassing all subject areas and works, to all client groups, this system will allow Griffith to target specific audience groups around the world.

## **eScience data archives**

There also exist a growing number of examples of eScience data repositories, which differ from ePrints in that they are designed to provide access to raw data files in various formats, as opposed to research papers. These services can also be hosted on repository systems. Griffith currently does not have an active central eScience data archive service project, although there exists potential for such a service using the repository.

Some issues include:

Purpose:

- It is unknown to what extent researchers will deposit and access such data collections;
- Would data owners want to use the datasets as working copies (to manipulate and run tests on) or as archive copies? This fundamental difference could greatly affect storage, management, usage;
- How will we measure, and what will be the return on investment for the institution?

Storage:

- Data collections are expected to be very large in file size;
- How will a digital repository store such large files, and what workflow processes are required to manage the files?;
- Data collections could be highly volatile. Some users may want to deposit data on a continual basis to share this with others.

Specialisation:

- Many and varying data formats requiring specialised services. The dataset is likely to require specialised structures, formats, and data query software or search scripts
- Providing such specialised services will require high level and continuous liaison with the researcher or academic as it is unlikely that central support staff will be expert in these areas.

## **Applying the digital repository to eScience**

Marketing efforts towards the ePrint and ePortfolio services will no doubt pave the way for user acceptance of an eScience service. Using an eScience data service as a purely archive service could require a different and possibly more simple service model than using an eScience service to host working copies of interactive datasets supporting current research. Collaboration can occur on a continuous basis during the research process and may result in superior final publications. Information Services could provide eScience services at cost to the client.

## **Conclusion**

Griffith University has been developing an integrated repository solution for the management of, and access to, digital collections supporting the learning and teaching and research goals of the University. The underlying repository structure has been established as have the bureau structures, metadata schemas and, workflows for learning objects, past exams and digitised course readings. An eprint service is also under development as is the research portfolio service. An opportunity exists for the development of an eScience archive service.

By customising workflows to the business needs of academic users, and by developing easy to use integrated interfaces, Griffith is maximising the opportunities to change academic user behaviour to achieve specific goals. Academics are being encouraged to share and reuse learning objects, and to select readings from library databases before requesting additional readings to be digitised. Academics will also be encouraged to deposit research publications into both ePrint and HERDC repositories and to leverage their research papers and data archive output using a research portfolio service. Further studies of user behaviour using server log statistics will be possible once these systems are in production. The desire to change user behaviour has been a major driver throughout the development of Griffith's digital repository project. By developing services within a single repository environment, Griffith is establishing a more sustainable and integrated service.

## References

- Bjork, B. 2004. Open access to scientific publications - an analysis of the barriers to change? *Information Research*, January 9 (2). <http://informationr.net/ir/9-2/infres92.html> (accessed 5 September, 2005).
- Borchert, M. and Richardson, J. 2005. Implementing the HarvestRoad Hive Digital Repository. *Information Online Conference*, February 2005, Sydney. <http://conferences.alia.org.au/online2005/papersandpresentations.php>
- Borchert, M. and Richardson, J. 2004. Integrating and digital repository system. *QUT Online Learning and Teaching Conference*, November 2004, Brisbane. <https://olt.qut.edu.au/udf/olt2004/index.cfm?fa=getFile&rNum=1586432&pNum=1333785>
- ePrints.org 2005. <http://www.eprints.org/> (accessed 1 September 2005)
- Joint Information Steering Committee (JISC), 2005. [http://www.jisc.ac.uk/index.cfm?name=project\\_cdlor](http://www.jisc.ac.uk/index.cfm?name=project_cdlor) accessed 5 September 2005 (accessed 1 September 2005)
- Malhotra, Y. and Galletta, D. 2004. Building systems that users want to use. *Communications of the ACM*, 70 (10), December 2004, 88-94.
- Markus, L. and Keil, M. If we build it, they will come: designing information systems that people want to use. *Sloan Management Review*. (Summer 1994), 11-25.
- Naijar, J., Ternier, S. and Duval, E. 2004. User behaviour in learning object repositories: an empirical analysis. *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1.
- Recker, M. and Pitkow, J. 1996. Predicting document access in large multimedia repositories. *ACM Transactions on Computer-Human Interaction*, 3 (4), 352-375.
- Uchitel, S., Kramer, J. and Magee, J., 2004. Incremental elaboration of scenario-based specifications and behavior models using implied scenarios. *ACM Transactions on Software Engineering and Methodology*, 13 (1), 37-85.