

ICE-RS

Table of Contents

1	Executive summary.....	2
1.1	Context.....	2
1.2	Deliverables.....	2
1.3	Benefits to the research community.....	2
2	Background.....	3
2.1	Terms used in this paper.....	4
3	Project objectives and strategies.....	6
3.1	Pilots.....	7
3.2	Extending the ICE system.....	7
3.3	Collaboration with APSR.....	7
3.4	Dashboard proof of concept.....	8
4	Intended Outcomes & timeline.....	8
5	Communication plan.....	9
6	Project structure.....	10
	Appendix A : Budget.....	11
	Appendix B: Biographies of Key Project Personnel	12
	Appendix C: References.....	13

1 Executive summary

This document is based on the proposal put to DEST for funding.

1.1 Context

The ICE-RS project seeks to improve national research effectiveness by providing infrastructure in the form of software tools, documentation and pre-packaged training

materials for academic authors, as part of the eFramework (Blinco 2006) for research.

This project complements, but is completely independent of, the proposed RAMP project by providing tools that address specific workflows related to document creation, repurposing and storage.

This is a sister project to one proposed by APSR to develop the Digital Scholar's Workbench, and the two teams will work very closely together on the shared parts of the systems.

Context

- Fits in to national research infrastructure
 - e-Framework
 - Complements RAMP
 - Related to APSR's 'Digital Scholar's Workbench'
- Dr Ian Barnes

1.2 Deliverables

ICE-RS will create software, manuals and training packages that allow and encourage efficient creation of flexible documents in the process of conducting and reporting on research, with benefits in improved efficiency, greater usability of research outputs and more sustainable research repositories.

Specifically, the project will enhance the Integrated Content Environment (ICE) which is a project of the University of Southern Queensland, with contributions from the RUBRIC project (RUBRIC 2005). ICE is a maturing system that is used within the university for courseware creation. The system gives a solid starting point for the ICE-RS project.

1.3 Contribution to systemic infrastructure

The FRODO round of projects are producing infrastructure for research, with products that provide services for research repositories through ADT, ARROW and APSR and federated security via the MAMS project.

MERRI projects cover a very broad range of infrastructure development some of which are looking at eResearch. None of the MERRI projects cover the end-to-end process of writing-up research using authoring tools and the associated publishing, annotation and review, document conversion and repository ingest processes.

ICE-RS is designed to deliver infrastructure that complements repository services by providing a field-tested open source product for writing research publications, with detailed documentation and training packages. This system is already published under an open source library, and will be extended as part of this project. The project will also provide connecting software to allow the system to be used in concert with FRODO-generated and recommended software for disseminating research.

The ICE-RS system involves adding two tiers of infrastructure below the Institutional

Repository, and will investigate the concept of a third 'dashboard' module to report on the state of both Institutional and working repositories. ICE-RS consists of:

New infrastructure
<ol style="list-style-type: none"> 1. A version-controlled, distributed 'working repository' for documents in progress. 2. An application for managing the process of writing research efficiently, for flexible delivery and long-term preservation. 3. A proof of concept reporting system, a 'dashboard' that will allow institutions to better manage their research activities by allowing dynamic reporting on work in progress.

Existing: Published repositories	Institutional systems containing finished written research outputs	3. “Dashboard” reporting comprehensive real-time statistics from both levels of repository.
1. Working repository	Research, and Courseware under development, with version controlled content that can be authored on or off campus and cross-platform.	
2. Staff workbench	Integrated authoring, backup, version control, collaboration, web publishing and reference management services (ICE).	

While the scope of work to develop these three components is a major undertaking, ICE-RS is able to build upon an established free software product, ICE, in which USQ has invested for its courseware requirements.

This contribution to Australia's systemic research infrastructure will be based on active trials with partners in the RUBRIC project, of which USQ is the lead, and targeted collaboration with FRODO projects.

Collaborations have been confirmed as follows.

Collaboration

ARROW

Integrate ICE-RS with the Vital repository.

APSR

Integrate ICE -RS with DSpace, share code and collaborate to make repositories more flexible.

RUBRIC partners

The following RUBRIC partners will participate in at least one pilot deployment of ICE-RS. USQ, USC, Massey.

e-JIST

e-Journal of Instructional Science and Technology- an international peer reviewed journal

A trial using ICE for submission and peer review will be undertaken.

JISC

JISC has expressed interest in assisting ICE-RS staff in undertaking a UK based trial. (Contact was made with Amber Thomas, Programme Manager, Information Environment).

1.4 Benefits to the research community

The benefits of the ICE-RS to national research effectiveness lie in:

Benefits to the research community

- Providing advice and freely available software to researchers for efficient authoring.
- Reducing institutional document processing costs through standardisation of work practices.
- Maximising the benefits of Institutional Repository projects by:
 - Making them able to deliver content more flexibly.
 - Improving sustainability by using standards-based document formats.
 - Reducing costs and increasing the amount of content captured by integrating repository ingest with research workflows.

USQ has already made a significant commitment to the ICE software, and will manage ICE-RS using existing project management resources and infrastructure.

The project will leverage existing university infrastructure in the same way as the RUBRIC project, drawing on services of the Distance and e-Learning Centre such as graphic design, project management, content management infrastructure, and IT support.

This project will be conducted in parallel with development work on the Scholar's Workbench System, produced by the APSR project.

2 Background

Two DEST-funded projects [APSR](#) and [RUBRIC](#), (RUBRIC 2005) along with the University of Southern Queensland's ICE (Integrated Content Environment) project have been collaborating on the development of systems for research-authoring with a view to maximizing the longevity of documents, increasing the flexibility of documents, and reducing document production, storage and publishing costs.

Currently in Australian universities a huge amount of effort is expended on writing documents in proprietary formats (such as Microsoft Word format) with little thought for the re-use or long term storage of these documents, or to the efficiency of authoring.

There are three main issues to be addressed

1. Ending the 'PDF only' nature of institutional repositories and expanding them to allow more flexible delivery.
2. Promoting document sustainability.
3. Increasing efficiency for researchers writing-up their research.

PDF-only nature of current systems

In the institutional repository world, the Adobe PDF format is currently the expected norm for document delivery.

Even though institutional repositories are web-based systems most content is not available in the native web format, HTML. HTML is more usable and flexible than PDF in many situations, allowing users to skim and sample content more easily than PDF. PDF, on the other hand, is a good solution for printing long documents and can be configured to make reading even book-length content a comfortable experience.

So why is it not the norm for repositories to offer both PDF and HTML?

It is because many of the widely used tools used for creating and storing research do not allow for reliable, automated production of HTML and PDF versions, and repository solutions are not geared to delivering content in flexible ways.

This proposal aims to continue developments already under way at ANU and USQ that have begun to break down the PDF-only nature of repository software and authoring processes.

(In technical fields, such as the maths and computing sciences, there are some solutions in use that can produce HTML and PDF output, such as the LaTeX typesetting system. These solutions can be integrated into the ICE-RS approach).

Sustainability

The Scholar's Workbench has addressed this concern by providing an easy-to-use system for authors to convert word processed documents into appropriate, XML-based documents. XML documents are both far easier to repurpose than traditional word processed documents, and are also suitable for long-term storage.

ICE addresses sustainability by using a profile of the OpenDocument standard (OASIS

2005). ICE-RS will be able to use the conversion code from the Scholar's Workbench to improve sustainability and provide more flexibility in document delivery.

Efficiency

Many researchers work in an ad-hoc way in word processing software supplied by their organization. The ICE approach can improve the efficiency of individual authors by providing templates along with a distributed version-controlled content management environment. This reduces effort in document formatting as well as in re-formatting for use in different environments.

ICE eliminates the need for hand-compiled tables of contents, gives every author access to PDF creation software and automatically produces HTML documents, reducing the need for re-processing of documents by web teams.

2.1 Terms used in this paper

Workbench

The term workbench appears in the name of the application developed at ANU, the Scholar's Workbench (Barnes 2006; Barnes 2005). The intention with this term is to capture the work-oriented perspective of researchers.

ICE is already functioning as a 'workbench' for courseware production; the ICE-RS project will expand that functionality into research-specific authoring.

Dashboard

In contrast to a workbench, which is a place for practitioners to work, a dashboard is a reporting mechanism allowing a university to introspect and to look at the state of work in progress in a way that enables decision making and planning.

The ICE-RS system will use a 'working-repository' which will be a centralized store of the research document authoring being undertaken across a university. The software to support this repository has already been developed as part of USQ's ICE project. ICE-RS will add metadata and information discovery features to aid in reporting and research management.

Each researcher will have access to their own documents. The centralized reporting system, dashboard, will be able to look into the document store and report on such things as:

- The number and type of documents under development.
- Target completion dates.
- Targeted publishers, journals and conferences.
- Research by subject.
- Completeness of research theses via metrics such as word-count and via ratings applied by supervisors.

These provide management indicators for IT support staff seeking to target their training resources, marketing departments looking for upcoming research that can be tied to a campaign, courseware authors seeking pre-release research, public relations seeking interesting opportunities, and the research office to aid in discovering nascent research, particularly where intellectual property might be at stake.

To enable this functionality, metadata will be collected as a document is created, rather after

the fact, which is the case with current institutional repositories.

This metadata collection process will take place in easy stages, with as much as possible supplied automatically by the system. For example ICE-RS will be able to automatically populate author and departmental affiliation metadata fields based on who has had input to a document. Other metadata, such as subject codes can be collected as a natural part of the authoring process, for example when the document moves from first draft stage to an internal review stage.

This metadata can then be used to ingest documents automatically without data re-entry.

Like the dashboard in a car, a scholarly dashboard would be a set of instruments that respond to the state of the machine, in this case an ICE -RS repository. Instruments would glean information directly from the research **as it is written**.

In this project we aim to develop proof of concept dashboard functionality that can provide some level of reporting.

Which area of dashboard functionality is chosen will depend on the nature of the pilot projects and the most pressing needs of the stakeholders. One might imagine that it would be useful to get a high-level view of thesis production, for example showing word-counts, the number of annotations made by supervisors, perhaps even a 'percentage complete' indicator across a number of theses.

This investigation into dashboard functionality is not intended to replace workflow systems or other research management systems but to complement them with real-time data gathered from work in progress.

Workbench and Dashboard functionality

Workbench

The ICE tool; with enhancements for research reporting.

Dashboard

Management views on content.

- How many words / citations / links?
- When is the paper / thesis / draft due?
- How many overdue conference papers in my department?
- How 'good' is my latest draft?

2.2 ICE

The Integrated Content Environment is a project of the University of Southern Queensland, developed to meet the university's requirements for a single-source print, web and CD-ROM courseware publishing system.

ICE is open source software. Version 1 is slated to become a supported university system in mid to late 2006. There are currently around twenty active users of the system. There are

expected to be hundreds of active users by mid 2007, using the system to write courseware, providing a base on which to build the ICE-RS extensions planned for this project. USQ has not promoted ICE for use outside of the university as it has not yet been released version 1.

The ICE template is an application profile for word processors, with two current manifestations. One for Microsoft Word and another for OpenDocument, most commonly used with the OpenOffice.org package.

ICE is of benefit to a very broad cross section of the national research community as it runs on multiple platforms, including Microsoft Windows, Apple Mac OS X and Linux / Unix.

ICE is already delivering benefits in the USQ context by simultaneously:

- Reducing document production costs by providing authors with proven templates and methodologies. Individual authors are more productive, and wasteful manual conversion steps are being eliminated or minimized.
- Increasing the quality and value of teaching documents through the use of intra-organizational standards for structure and presentation.

Even without the additional funding sought, ICE is beginning to make inroads into the Research arena, with a modest pilot under development in the Australian Centre for Sustainable Catchments. However, development resources are limited and the funding sought is needed to undertake the significant enhancements that the ICE software requires to be used in a research context.

The following is the feature set of ICE as described in a paper on the system (Sefton 2006) to be delivered at AUSWEB 2006:

The requirements for ICE were that the system:

1. Run on Windows, Mac OS X and Linux.
2. Be usable on and off campus and from home, with or without a network connection.
3. Work in both Microsoft Word and OpenOffice.org Writer, using a generic template containing a set of pre-defined styles.
4. Be usable by non-technical and non-specialist staff.
5. Produce book-length course content in PDF for traditional distance education use.
6. Produce web-optimized content in HTML and PDF from the same collection of source documents as the print version.
7. Produce IMS content package format for use in Learning Management Systems.
8. Maintain a full revision history across multiple instances of a course and deal with courses offered in multiple ways with some changes.

2.3 Digital Scholar's Workbench

The Digital Scholars' Workbench (DSWB) is a prototype tool that converts well-structured word processing documents to a usefully-deep XML structure. The DSWB acts as a middleware between the standard word processors and a rich set of mature, XML-aware web publishing, print publishing, and workflow applications.

3 Project objectives and strategies

The high-level aims of the project are to produce software that adds to the infrastructure and know-how created by other FRODO and MERRI projects:

- a. To conduct pilots using the ICE content creation system to write research theses, papers and books.
- b. To add features to the ICE software to aid in research workflows, chiefly indexing of full-text and document metadata.
- c. To collaborate with the APSR project on the joint development of templates, software and documentation wherever possible.
- d. Demonstrate a proof of concept 'dashboard' for universities and research centres able to extract metadata from research as it is being documented, and display management indicators. This would be done in harmony with the RAMP project's broader scope.

3.1 Pilots

A range of other institutions, including three RUBRIC partners, the e-Framework project and JISC have agreed to collaborate on at least one of the following pilots, taking one or more pieces of research from authoring to repository ingest using ICE-RS and appropriate repository integration.

Establish pilots at USQ and across the RUBRIC partner network:

- a. At least five each of honours, masters and PhD theses, with infrastructure enabling them to go from authoring to examination to ingest into the institutional repository.
- b. At least ten research articles for peer review, including both journal and conference papers, and associated presentations.
- c. At least two books.

3.2 Extending the ICE system

Establish pilots at USQ and across the RUBRIC partner network:

- a. Produce 'how to' guides for writing-up research using the ICE template.
- b. Develop proof-of-concept ICE template to article / book formatters for one or more publishers or conferences.
- c. Develop proof-of-concept repository ingest for Fez / DSpace / Vital (the recommended Institutional Repository software coming out of the FRODO projects) for ICE documents.
ARROW and APSR have indicated that will collaborate with ICE-RS to program repository integration.
- d. Integrate with MAMS security projects, and RAMP to allow fine-grained control over access to resources and parts of resources.
- e. Develop a set of guidelines for using bibliographic software with the ICE template.
- f. Add full-text indexing with title and subject indexed for resource discovery.
- g. Add an annotation / voting / approval system for collaborative editorial review processes (not peer review in the formal sense)
Allow documents to be labeled with states such as 'first draft' or 'submitted for examination' so that dashboard reporting systems can be used to summarize the state of a set of working documents.

3.3 Collaboration with APSR

- a. Ensure that documents are inter-operable between ICE and DSWB.
- b. Build a 'Deposit this' button for ICE to DSpace ingest.
- c. Meet with the APSR team on a regular basis.

3.4 Collaboration with ARROW

- a. Build an ICE to VITAL single-click 'Deposit this' button.
- b. Contact ARROW partners with a view to conducting ICE-RS pilots.

3.5 Dashboard proof of concept

- a. Demonstrate granular indexing of content within documents, for example being able to search research goals and compile summary documents.
- b. Develop demonstration 'Dashboard' reporting for university service departments, based on a needs analysis.

4 Intended outcomes & timeline

The intended outcomes of the project are that the existing ICE project is augmented. ICE will have:

1. New features supporting research which will have been developed in parallel with live pilots and intense stakeholder involvement.
2. Guides and training materials to enable universities to add ICE to their research workflows.
3. A demonstration of institutional 'dashboard' functionality, allowing reporting on research activities from the corpus of research documentation, as well as via RAMP software.

Specific activities and deliverables by quarter are listed below, the intention is to integrate RAMP deliverables as they become available.

2006 Q4 – 2007 Q1

1. Recruit staff.
2. Train staff in ICE.
3. Recruit pilot users and train them in ICE.
4. Add ICE configuration and features to support new document types for research.
5. Hold introductory workshops with pilot users.
6. Scope work needed to integrate with RAMS.

2007 Q2

1. Implement full-text indexing and metadata in ICE
2. Support pilots.
3. Begin augmenting ICE documentation and training materials.
4. Add features for RAMS integration as appropriate.
5. Begin configuration for conversion to required journal, conference and book formats.
6. Scope Repository Ingest.
7. Hold workshops with pilot users to determine further development priorities.

2007 Q3

1. Develop an institutional dashboard proof of concept to report on the state of research documentation held in ICE.
2. Support pilots.
3. Continue configuration for conversion to required journal, conference and book formats.
4. Scope Repository Ingest.
5. Freeze new features for the next quarter.

2007 Q4

1. Finalize pilots or organize for ongoing support post ICE-RS project.
2. Publish final documentation.
3. Produce final software release.
4. Report on project outcomes at an event to be held in conjunction with RUBRIC.

5 Communication plan

The dissemination of ICE-RS project outcomes will be facilitated through a number of specific communication mechanisms, including:

- An ICE-RS website will be established which will be hosted and maintained on the USQ Content Management System. Public project documentation will be made available, together with links to relevant sites.
- A discussion list will be established for interested parties to provide a continuous dialog mechanism on project developments.
- A regular ICE-RS project newsletter will be distributed.
- A weblog will be established to disseminate information regarding project progress that can be accessed by the RSS aggregators of interested parties.
- Selected stakeholders will be provided with access to the ICE-RS Trac Project management website (which includes a wiki) for the open and accountable monitoring of project milestones and tasks.
- Publications will be submitted to relevant conferences and journals.
- Conference presentations will be delivered at significant national forums such as Open Publish, Open Repositories Forum, and EDUCAUSE enabling the distribution of key project outcomes to the wider higher education sector.
- Meetings will be held with associated projects to bring together representatives for complimentary strategic planning and regular updates on activity.
- A major ICE-RS event will be held to demonstrate project outcomes to the wider higher education community. There are also opportunities for joint activities with the RUBRIC and APSR projects to leverage off existing infrastructure and share costs.
- Marketing to the research community will be done via a series of campaigns primarily aimed at Research Centre and Faculty researchers and administrators. There will also be

an opportunity access the RUBRIC partner institutions, to raise awareness of repository ingest and document sustainability issues.

6 Project structure

The ICE project is an established project with:

- 2.5 FTE staff handling development and roll out.
- A suite of project management tools and an established, customer focused development process.

The funds sought here would supplement the team with an additional resources for one year:

1. 1 FTE project officer to organize pilots, liaise with university staff and students, and produce documentation for ICE in a research context.
2. 1 FTE technical officer (programmer, systems administrator) to add features to the ICE program, chiefly full-text indexing and standards-compliant metadata handling along with formatters for conference and journal formats.
3. Travel for partners participating through the RUBRIC partner network and for collaboration and harmonization with APSR.

Advice for the project team will be sourced from:

- Direct involvement in the ICE process from key USQ personnel, Peter Sefton and Alan Smith.
- The RUBRIC project and its board.
- Close links with a sister project from APSR and USQ involvement in the RAMP project.

References

Barnes, I. 2005. Sustainability of word processing documents.
<http://www.apsr.edu.au/currentprojects/wordprocessing.html> .

———. 2006. Integrating the Repository with Academic Workflow. *OpenRepositories*.
http://www.apsr.edu.au/Open_Repositories_2006/ian_barnes.pdf .

Blinco, K. 2006. The JISC -DEST eFramework - integrating everything. *Open Repositories*.
<http://www.e-framework.org/> .

OASIS. 2005. OpenDocument v1.0 specification. <http://www.oasis-open.org/committees/download.php/12572/OpenDocument-v1.0-os.pdf> .

RUBRIC. 2005. RUBRIC Grant Submission Intended Outcomes.
<http://rubric.edu.au/bid/bidoutcome.htm> .

Sefton, P. 2006. The integrated content environment.