

JISC DEVELOPMENT PROGRAMMES

Project Document Cover Sheet

COMPLETION REPORT

Project

Project Acronym		Project ID	
Project Title	Theses Alive!		
Start Date	Oct 2002	End Date	Oct 2004
Lead Institution	The University of Edinburgh		
Project Director	John MacColl		
Project Manager & contact details	Theo Andrew Edinburgh University Library George Square, Edinburgh EH8 9LJ		
Partner Institutions			
Project Web URL	http://www.thesesalive.ac.uk/		
Programme Name (and number)	FAIR		
Programme Manager	Rachel Bruce/ Balviar Notay		

Document

Document Title	Completion Report		
Reporting Period	n/a		
Author(s) & project role	Theo Andrew (Project Officer) John MacColl (Project Director)		
Date		Filename	
URL	<i>if document is posted on project web site</i>		
Access	<input type="checkbox"/> Project and JISC internal	<input type="checkbox"/> General dissemination	

Document History

Version	Date	Comments
0.1	Oct 2004	First draft
0.2	Jan 2005	Second version with budgetary details.

Project Overview

1. Background

At its most mundane an electronic thesis is a digital image of the print distribution object, being still firmly grounded in the traditions of print. However, the digital format offers a unique opportunity to create an electronic document unrestricted by conventional limitations. It is now possible to author a document that contains both aspects of multimedia and the dynamic presentation of large data sets that previously were unattainable in print format. In addition to these benefits, the digital format takes full advantage of the networked computing environment to deliver the thesis literature, which has for too long been considered intractable, online to a global audience.

The practice of making theses and dissertations available online is growing internationally. Repositories of electronic theses and dissertations are now common in universities in North America, Australia and in many European countries. However, the UK has generally been relatively slow in adopting electronic theses, but a number of recent initiatives have tried to promote the electronic thesis agenda within the UK HE context. The foremost of these initiatives was the UK Theses Online Group (UTOG).

In October 2001 the SELLIC team at Edinburgh University Library (EUL) presented a report to the UK Theses Online Group on the results of our Edinburgh University Library Doctoral Theses Digitisation Project. This work was generously funded by UTOG. This report contained certain recommendations for the future of the management of 'born digital' theses at Edinburgh. It concluded that universities were moving into a digitally networked environment which has the potential to transform the current system for providing access to theses, by making them freely available on the web.

The Theses Alive! project was originally conceptualised as part of the JISC-FAIR programme to consolidate the preliminary work on electronic theses and to advance the understanding of processes and procedures and technical infrastructure required for electronic theses to be adopted by the UK HE community. The project has now delivered on this conceptual aim by building an exemplary electronic thesis repository, whilst in the process determining the best practices involved, and providing the tools necessary for other HE institutions to follow.

2. Aims and Objectives

The Theses Alive! objectives as stated in the initial project plan were:

1. To develop a thesis submission system for use in all participating universities
2. To develop and support a generic metadata format capable of delivering metadata to a number of relevant metadata repositories for UK thesis information
3. To develop an infrastructure which enables e-theses to be published on the web to the extent that a minimum of 500 e-theses exist within the UK segment of the NDLTD after two years
4. To test the value of a national support service for e-theses creation and management in the UK
5. To work with other e-theses developments internationally, and in particular to assist the research aims of other e-theses projects funded within the JISC FAIR Programme.
6. To produce a 'checklist approach' for universities to use as they develop e-theses capability.

As the project progressed it became apparent that a number of these objectives did not reflect the developing needs of the project. Although it was not explicitly expressed in the original project plan, we found it necessary to investigate a number of issues, such as intellectual property rights associated with electronic theses, and implications of the impending Freedom of Information Act. These were not budgeted for in the original plan; however the flexible approach adopted by project staff allowed these issues to be addressed. Overall, the Theses Alive! project met most of the objectives, with the exception of objective 3. Even though the infrastructure was developed, the project failed to accumulate 500 electronic theses. Edinburgh archived 189 theses, and our estimate of the total number of e-theses in the UK segment of the NDLTD is approximately 230. With the benefit of hindsight, the figure of 500 was perhaps optimistic at best, being based on informal co-operation with five other institutions.

3. Overall Approach

Delays in setting up the Project due to summer vacation and the reallocation of existing contract-based staff into the Project led to us to reschedule the Project to begin in November 2002. Subsequently the departure of the original Project Officer at the end of January resulted in further delays. A replacement was not in post until May 2003; however, these delays did not affect the overall project in an adverse way.

One of the major drawbacks of the project was the informal nature with the project partner institutions. Section 11 (Project Resources: Project Partners) details the major problems with this arrangement. With the benefit of hindsight, if the project was to be re-run, each institution should formally agree their level of commitment and involvement prior to the project starting. We would also wish all partners to receive some funding in order to secure their commitment.

4. Project Outputs

All major project outputs, discussed in the following sections, have been submitted to JISC where appropriate.

5. Project Outcomes

The main achievements and key findings of the project are as follows:

1. We have developed best practice guidelines for institutions wishing to adopt electronic theses, which are reflected in the project documentation.
2. We have investigated and are in the process of implementing revised thesis rules and regulations for the University of Edinburgh to permit submission of electronic theses and dissertations.
3. Similarly we have updated the thesis submission and management workflow process to take advantage of the benefits that adopting electronic theses create.
4. We have delivered a report on IPR and electronic theses which was commissioned by JISC Legal. This stemmed directly from work carried out by the project.
5. Sample use and deposit licences have been developed, including advice on the current topic of Freedom of Information implications.
6. Community support has been achieved via project documentation and journal publications.
7. We have actively helped to assess and explore different mechanisms for the disclosure and sharing of content, by evaluating open source software platforms for digital repositories (e.g. Jones, 2003) through to assessing the cultural use and impact of digital media (e.g. Andrew, 2003).
8. Creation of the recommended UK e-thesis core metadata set, alongside the Robert Gordon University, the University of Glasgow and the British Library.
9. Development of the Thesis Alive Plug-in for Institutional Repositories (Tapir), adding e-thesis functionality to the internationally used DSpace software platform.
10. Launch of the Edinburgh Research Archive (ERA) - University of Edinburgh's institutional repository.
11. The pilot electronic thesis project in two academic schools in the University of Edinburgh gathered c.170 electronic theses and delivered them online.

These project achievements directly confer the following dynamic impacts on the teaching, learning, and research communities:

1. The technical and cultural expertise we have garnered through developing and implementing Edinburgh's live institutional repository service can and will be disseminated to the HE information and library services community. The hard-won lessons we have learned will make this process for other institutions a much more enriched one.
2. The major impact that this project has delivered to the research community is the provision of Open Access to selected research and thesis literature from the University of Edinburgh.
3. This has a knock on effect for enhancing teaching and learning, in that source material, e.g. book chapters etc, are also increasingly being made available through the repository. This toll-free access to students is available constantly without the physical lending restrictions that are traditionally associated with published literature.

4. Most importantly, the internal thesis submission and subsequent management processes at Edinburgh have been updated enhancing the student teaching and learning environment.

In conclusion, the following points learned during the project lifetime are directly applicable elsewhere:

1. Experience in advocating culture change through the promotion of Open Access, and the acceptance and use of digital documents in scientific and scholarly communication.
2. Expertise in setting up and managing an exemplary institutional repository using DSpace.
3. Acquiring a detailed knowledge of how to lead software development in an open source environment.

A number of exciting opportunities and outcomes have arisen from experience gained during the project:

1. Primarily, the Theses Alive! systems developer, as a result of work carried out during the project, has become highly regarded internationally in the open source software community such that he has been given the prestigious role of one of the team of 'committers' for the DSpace federation, responsible for the future development of the codebase. Only a handful of trusted developers are given this access.
2. Similarly, the work this project has achieved in developing DSpace for e-theses has been recognised by the Networked Digital Library of Theses and Dissertations¹ (NDLTD), which has elected us to be inaugural members of the ETD DSpace Implementers Group².

A one page summary describing the project achievements is attached as Appendix A.

6. Stakeholders

The obvious stakeholder beneficiaries are postgraduate students, who will benefit from an integrated thesis submission system that will speed up the whole process by cutting out duplication of labour, and by reducing their overall workload by centralising the printing/binding process from the copy supplied electronically. Additionally, students will benefit through the proven additional readership that electronic theses facilitate, and the ability to easily reference new and old thesis literature. Similarly, other researchers worldwide are likely to benefit from the provision of open access research theses by being able to access a previously restricted, yet valuable, literature.

The second major stakeholder group that will directly benefit from the project outcomes is the Library and Information Services community, and other Open Access advocates who are responsible for the development and implementation of institutional repositories. The work and experience this project has acquired in developing the technical infrastructure, and appropriate policies, should be used by other interested parties to help foster their own institutional repository progress.

7. Intellectual Property Rights

Where applicable all of the project outputs have been circulated to the teaching, learning, and research communities under open-source licences. The most significant Theses Alive! product which will outlive the project lifetime is the Tapir (Theses Alive! Plug-in for Institutional Repositories) software code. This software has specifically been issued with a BSD-style (Berkley Source Distribution) licence, attached as Appendix B. This licence specifically allows freedom for users to redistribute and use the source and binary forms, with or without modification, whilst offering adequate legal protection for the University of Edinburgh against liability.

A number of outstanding IPR issues exist, the most significant of which is the use of third party copyright material in theses. Traditionally, with unpublished theses and dissertations this was not a problem as the fair dealing defence permitted this use. However, hosting material online in any legal sense is considered publishing, thus any copyright material contained within e-theses must be cleared for use before being archived. As there is no legal precedent for this problem it is hard to predict any future outcome.

¹ NDLTD:<http://www.ndltd.org/>

² ETD DSpace Implementers Group: <http://ivlab3.ils.unc.edu/~jfox/etd-dspace/>

Project Resources

8. Project Partners

The Theses Alive! project initially liaised with five other HE institutions, consisting of The University of Cambridge, Cranfield University, Leeds University and Manchester Metropolitan University. The initial project concept was to develop repository software to be used by these universities, and to test a national support service for this group of pilot institutions. Being unfunded these institutions did not form a formal consortium, and it was understood from the outset that any participation was purely on a voluntary basis. But due to the rapidly changing environment for repository software in the UK over the past two years, it became clear that the project agenda did not fit with all of the individual institutional aims and objectives. Given the project timescale it was necessary to progress with the project's core development without these pilot institutions fully on-board to test the end products.

Similarly it became apparent that a national e-theses support service was not entirely appropriate. Although it is necessary to help institutions build repositories and appropriate policies, it was felt that other types of support, for example student support or mediated deposit, would be best offered by the home institution where local staff would have detailed knowledge of current working practices and procedures.

Although there was no tension between the Edinburgh team and any of our partners, we have drawn the lesson that when a project is working with partner institutions, the responsibilities, aims and objectives of each institution should be made clear at the start of the project and formalised in some capacity, otherwise the final project outcomes may suffer. It is very difficult to maintain the commitment of either the partners or the core team when there is no funding involved. The partnership work is done on a 'best endeavour' basis and deprioritised when other deadlines are pressing. It is important that, if unfunded partners are deployed, these factors are understood clearly from the beginning.

The Theses Alive! project did foster close links with other projects within the FAIR programme, specifically those within the eFAIR cluster, and the e-thesis sub-cluster. Working with the Electronic Theses project (Robert Gordon University) and the Daedalus project (Glasgow University), the sub-cluster developed a recommended metadata core set for e-theses. Working within this cluster environment proved to be extremely beneficial, providing regular contact and support (see section 13).

9. Project Management

Theses Alive! was a relatively small project, funded over just two years, with a staffing complement of only 1.35 FTE throughout that period. Just over 85% of the total award from JISC was spent on staff. It was always intended that Theses Alive! would be a practical project, intended to produce a real working system to serve as an exemplar for consideration across the UK. Given that the objectives we set ourselves were challenging, we took a decision at the beginning of the project to focus on achieving strong deliverables using a tightly-knit team. We believe that this has been a successful strategy, and that the basis of the project within a single institution has had a major part to play in this. It has allowed team members to communicate easily, face-to-face, and this has eliminated much of the cost of communication, both financial and in time lost, which is common in projects of this type. The disadvantage has perhaps been too much of a reliance upon the core team for delivery of the project goals, and the volunteer partner sites have not been allowed to contribute as much as they might have. Since the project was small in personnel terms, we also took the risk of sharing the project management between the Director (0.1 FTE) and the Project Officer (0.5 FTE), a decision which would not have worked in the case of a larger project with several funded partners.

Nevertheless, while the project did concentrate its development effort within Edinburgh, we believe that we contributed strategically within the FAIR Programme as a whole. Team members were present at all of the FAIR cluster meetings, and we had a high profile in the e-theses 'sub-cluster', working alongside our colleagues in Glasgow, Robert Gordon's, Cranfield and the British Library. Indeed, this wider professional involvement became very necessary as a means of compensating for the close nature of the team at Edinburgh. It was important both to test out our ideas and our developing system on colleagues working on related projects within the FAIR Programme. The team regularly met up with these colleagues not just at JISC-run meetings, but at other events within the UK and abroad, and this permitted a considerable amount of cross-referencing of each other's work at

conferences. The shared effort on the UK E-Thesis Core Metadata Set demonstrated the ability of the team to interwork and collaborate successfully across the Programme. The Project Officer and the Systems Developer were also involved in Edinburgh's contribution to the SHERPA Project – a factor which became particularly useful once it was decided to use DSpace as the platform for both our e-theses and our eprints here at Edinburgh.

Clearly, the small-scale, practically-focussed and development-driven model which Theses Alive! represents will not be appropriate to all projects, but we would recommend that it be encouraged where possible. One of the key success factors has been our ability to use our own institution as a 'guinea pig' for our development efforts, and it has been rewarding to see that the work we have done in building a repository for the management of e-theses and eprints has been warmly welcomed at Edinburgh, and that the repository will become a central university service. The need to demonstrate the value of the system meant that the project staff spent a lot of time in communication with library, ICT and academic colleagues, and as a consequence they came to be viewed as University staff first, and project staff second. This has had the advantage both of allowing them to focus on the real developmental, liaison and marketing needs (since what they were building was wanted by real users by real deadlines) and of making themselves sufficiently important to the University that they have been able to continue in employment within us for at least eight months beyond the project end-date, on library funding.

The personnel at the end of the Project were:

John MacColl, Director (0.1 FTE)
Richard Jones, System Developer (0.75 FTE)
Theo Andrew, Project Officer (0.5 FTE)

John MacColl has gained experience in managing a small, focussed development project, and in understanding the DSpace development community. Among other benefits, this has helped with his expertise in institutional repositories more generally, and has led to requests to speak about Edinburgh's DSpace experiences in various contexts.

Richard Jones has developed considerable skills in Java programming, in metadata and in technical project management. He has also gained experience in writing articles for publication, technical and end-user documentation and in making presentations, and has very successfully learned how to convey technical information to non-technical audiences. His particular experience in writing DSpace code has made him an important member of the DSpace Code Committers' Group.

Theo Andrew has developed from joining the project as a newly-completed PhD student himself, into an enthusiastic information scientist with a particular interest in digital research repositories and in intellectual property in relation to research output. He has also gained considerable experience in liaising with the academic community, in marketing new services and in making presentations. He now plans to embark upon a course of study leading to an appropriate qualification in information science.

10. Programme Support

Initially the project benefited greatly from the support of a proactive programme manager, and an active calendar of training and support events. The *ad hoc* arranged workshops, e.g. the IPR and sustainability workshops, worked especially well and allowed the various extra-curricular needs of the projects to be addressed.

Similarly the FAIR programme was set up in such a way that regular contact with other projects was encouraged. This innovative organisation of projects into clusters, and where appropriate sub-clusters, allowed the FAIR projects to support each other, via regular contact. This encouraged close collaboration to achieve common aims and objectives. Such cooperation between the Edinburgh-led Theses Alive! project, the RGU-led Electronic Theses project and the Glasgow-led Daedalus project facilitated the creation and adoption of the UK E-Theses Core Metadata Set.

11. Budget

The budget is shown in Appendix B. A negative balance of -£8,155 is shown at the end of the Project. This resulted from a higher expenditure on travel than anticipated (due mainly to international travel to NDLTD conferences). The additional costs were borne by Edinburgh University Library. No funds were obtained from other sources.

Detailed Project Planning

12. Evaluation Plan

Quantitative evaluation has occurred on three main levels, corresponding to the main project outputs;

<i>Evaluation area</i>	<i>Evidence</i>	<i>Measurement</i>
Support	Use of Project documentation	Download statistics from project website.
Technical	Third party use of the Tapir software	Download statistics from source code repository
Institutional Repository	<p>Researchers and students using the repository.</p> <p>External researchers using the repository.</p>	<ul style="list-style-type: none"> - Number of items deposited in ERA. - Number of searches carried out, - Number of items downloaded from ERA.

Table1: project evaluation areas and criteria.

12.1 Support evaluation

The most effective way to measure the impact of the support we offered to other institutions is to look at the usage statistics from the project documentation archived on the Theses Alive! website. Overall, between 12/16/2002 to 11/1/2004, 4,793 unique users came to the site, as determined by typical behaviour of browsers with a non-rotating IP address and including a projection of the true number of visitors with rotating IP addresses. These users visited the web site on 16,556 occasions, with each typical visitor examining 5.17 distinct files before leaving the site. A typical visit lasted for 7.87 minutes.

Appendix C shows the top 40 downloaded items from the Theses Alive! site archive, which contains much of the output of The Theses Alive! project. This includes the original project proposal, and subsequent alteration and progress reports. In addition there are self-archived publications as well as reports which are of interest, but not for publication due either to their size or scope. All events attended by the Theses Alive! team are also written up and placed on these pages. Similarly all conference powerpoint presentations have been archived. Finally, documents concerning the creation of the EUL-DSpace product and DSpace itself can be found here. In total, these top 40 documents have cumulatively been downloaded 10,827 times. This would suggest that the documentation we have made available through the Theses Alive! website has been extremely well-received and is a valuable contribution to the HE library and information services community.

12.2 Technical evaluation

The Tapir was developed to add e-thesis functionality to DSpace, one of the most popular institutional repository software packages available internationally. To assess the impact that the Tapir has made

it is possible to analyse the statistics from the independent source code repository where it is stored (<https://sourceforge.net/projects/tapir-eul/>).

Appendix D shows the Tapir download statistics in detail. In total, the Tapir has been downloaded 31 times. It is worth noting that this software is highly specialised and would only be of interest to those wishing to implement an e-thesis repository using DSpace. Approximately 120 institutions have working DSpace repositories worldwide, thus the Tapir download figure suggests a significant proportion of those institutions are interested in our developments. It is interesting to see that the Tapir has generated a lot of worldwide interest, with institutions such as Bergen University and the University of New Brunswick downloading and installing the Tapir for use in their own repositories. It is clear that institutions worldwide have found the Theses Alive! project outputs to be an extremely valuable contribution.

12.3 Repository evaluation

One of the major outputs of the project is the development of the Edinburgh Research Archive. To evaluate this it would be useful to assess how ERA has been received by researchers at the University of Edinburgh, and also by researchers worldwide. Perhaps the most comprehensive way of achieving this is to look at the cumulative statistics from the server and web logs. These give a quantitative view of usage, in terms of items submitted, items downloaded and the number of searches carried out in the archive database, as shown in Appendix E.

A simple look at the statistics shows that, during the project lifetime, 189 electronic theses and dissertations were submitted to ERA. Even though this total number of theses is overall smaller than intended, and leaves the UK as a whole some way short of the 500 e-theses we hoped for at the outset, it still represents a significant effort, and a sufficient proof-of-concept demonstration of the demand for and operational viability of e-theses repositories in the UK.

13. Quality Assurance Plan

The following QA standards have been followed for the main technical developments that the project has undertaken.

13.1 Website Accessibility

Both the Theses Alive! and ERA websites, to the best of our knowledge, conform to the University of Edinburgh's guidelines for accessibility. The design continues to be enhanced, as far as possible within resources, in order to improve accessibility and usability. Templates are also tested for reasonable backward compatibility. Specifically:

- Font sizes are relative so users may resize text according to needs
- Abbreviation and acronym tags are being added
- Stylesheets are being used extensively to separate content from presentation
- Content should still be visible to browsers with no support for stylesheets or Javascript
- Pages are resolution-independent and able to expand or contract to fit users' screens
- All images used to convey any information of any kind have alternate text (alt) tags
- Graphics used purely in the construction of the page design or layout have empty alt tags.
- As far as possible, links are written to make sense out of context
- Main links have titles added which display in many visual browsers as "tooltips" but are also used by assistive technologies
- Where appropriate, text versions of pages are provided
- Templates have been checked and the front page adjusted for users with High Contrast settings

All of this is constantly under review but, in addition, we aim to achieve close conformity with web accessibility guidelines suggested by the W3C Web Accessibility Initiative³.

³ <http://www.w3.org/WAI/WCAG1A-Conformance>

13.2 Repository standards

The Edinburgh Research Archive operates with the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH v.2). This interoperability standard means that metadata within our database is harvested by service providers, who can aggregate the data to provide federated searches across independent repositories.

14. Dissemination Plan

14.1 Journal Articles

1. MacColl, J. (2002), 'Electronic Theses and Dissertations: a Strategy for the UK', Ariadne, Issue 32 July 2002

<http://www.ariadne.ac.uk/issue32/theses-dissertations/>
http://www.thesesalive.ac.uk/arch_publications.shtml

2. Andrew, T. (2003), 'Trends in Self-Posting of Research Material Online by Academic Staff', Ariadne, Issue 37 October 2003

<http://www.ariadne.ac.uk/issue37/andrew/>
http://www.thesesalive.ac.uk/arch_publications.shtml

3. Jones, R. D. (2004), 'DSpace vs. ETD-db: Choosing software to manage electronic theses and dissertations', Ariadne, Issue 38 January 2004

<http://www.ariadne.ac.uk/issue38/jones/>
http://www.thesesalive.ac.uk/arch_publications.shtml

4. Andrew, T. (2004), 'Theses Alive! : an E-theses management system for the UK', Assign, Volume 21, Issue 3 April 2004

http://www.thesesalive.ac.uk/arch_publications.shtml

5. Andrew, T. (2004), 'Intellectual Property and Electronic Theses', JISC Legal Publications, September 2004

<http://www.jisclegal.ac.uk/publications/ethesesandrew.htm>
http://www.thesesalive.ac.uk/arch_publications.shtml

6. Jones, R. D. (2004), 'The Tapir: Adding E-Theses Functionality to DSpace', Ariadne, Issue 38 November 2004-11-02

<http://www.ariadne.ac.uk/issue41/jones/>
http://www.thesesalive.ac.uk/arch_publications.shtml

14.2 Conference Presentations

1. *An ETD Submission System for the UK*. John MacColl. Sixth International Symposium on Electronic Theses and Dissertations, 2003, May, 20-24. <http://edoc.hu-berlin.de/etd2003/maccoll-john/>

2. *Theses Alive! - an ETD Management System for the UK*. Theo Andrew, Richard Jones OA Forum Workshop, Bath. 5th September 2003 http://www.thesesalive.ac.uk/arch_confpapers.shtml

3. *ERA: The Edinburgh Research Archive*. Theo Andrew Capturing Edinburgh's Research, Edinburgh. 23rd October 2003 http://www.thesesalive.ac.uk/arch_confpapers.shtml

4. *Electronic Theses: The Story So Far*. Theo Andrew Scholarly Communication Seminar, Leeds. 30th October 2003 http://www.thesesalive.ac.uk/arch_confpapers.shtml

5. *Theses Alive!* John MacColl E-Theses: 'A Future for UK Theses' Seminar, London. 22nd January 2004 <http://www2.rgu.ac.uk/library/uk-theses.htm>

6. *Technical Issues for Repository Software*. Richard Jones
ePrints UK Workshop, Oxford. 22nd March 2004
<http://www.rdn.ac.uk/projects/eprints-uk/workshops/oxford/>
http://www.thesesalive.ac.uk/arch_confpapers.shtml

7. *Why self-archive? What's in it for institutions and for libraries?*. Theo Andrew.
ePrints UK Workshop, Edinburgh. 14th May 2004
<http://www.rdn.ac.uk/projects/eprints-uk/workshops/edinburgh/>
http://www.thesesalive.ac.uk/arch_confpapers.shtml

8. *Technical Issues for Repository Software*. Richard Jones
ePrints UK Workshop, Edinburgh. 14th May 2004
<http://www.rdn.ac.uk/projects/eprints-uk/workshops/edinburgh/>
http://www.thesesalive.ac.uk/arch_confpapers.shtml

9. *DSpace for E-Theses at the University of Edinburgh and Robert Gordon University*
Richard Jones, Andrew Penman (RGU). ETD 2004, Lexington, Kentucky. 3rd June 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

10. *ERA: Capturing Edinburgh's Research*. Theo Andrew
ETD 2004, Lexington, Kentucky. 4th June 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

11. *DSpace for E-Theses*. Richard Jones
E-Theses: The Next Step, British Library, London. 27th September 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

12. *Intellectual Property and E-Theses: What you need to know*. Theo Andrew
E-Theses: The Next Step, British Library, London. 27th September 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

13. *DSpace Ingest Workflow*. Richard Jones
ERPANET Workshop on Workflow in Preservation, Central European University, Budapest
13-15th October 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

14. *DSpace at Edinburgh*. Theo Andrew
SPARC Institutional Repository Workshop
Washington DC. 18-19th November 2004
http://www.thesesalive.ac.uk/arch_confpapers.shtml

15. Exit Plan

With the submission of this document, all core project documents have been submitted to JISC.

All core project documentation and reports have been made available through the project web site (<http://www.thesesalive.ac.uk/>). Similarly all important project outputs will be deposited in Edinburgh's institutional repository; the Edinburgh Research Archive. All software developed is Open Source and maintained and hosted at SourceForge.net; the world's largest Open Source software development web site.

As specified under the terms agreed with JISC in the letter of grant and the JISC Terms and Conditions attached to it, the University of Edinburgh will continue to host the project web site for three years after project completion, and will assist JISC in archiving it subsequently.

16. Sustainability Plan

The major project output that will be sustained after the project deadline is the electronic thesis repository. Early on in the development it was decided that it would be easier to sustain if it was integrated with an Edinburgh e-prints repository, which was simultaneously being developed as part

of the SHERPA project. Thus, both major project outputs converged to create the Edinburgh Research Archive. The Edinburgh Research Archive is seen as a core strategic service offered by Edinburgh University Library, which is represented by its prominent position in the University's Knowledge Management Strategy.

Many of the projects outcomes are being followed up in the forthcoming JISC-funded EThOS project, however the relationship between institutional repositories and the impending Research Assessment Exercise needs to be investigated further, and if appropriate, technical systems based on existing IR platforms (DSpace and EPrints) need to be developed to add the required functionality.

Appendix A. Summary of Project Achievements

The *Theses Alive!* project has developed on a pilot basis a distributed system for the management of electronic theses and dissertations (ETDs) in the UK, in order to take advantage of the availability of the networked digital environment. The system we developed provides both a metadata submission system, and a full-text online repository for those universities wishing to use it. To achieve these aims we initially identified six objectives that the project needed to focus on to be successful; i) Develop a digital thesis submission system for use by interested universities, ii) Develop an international standards compliant digital infrastructure which enables e-theses to be published online, iii) Develop and support a generic metadata format capable of delivering metadata to a number of relevant metadata repositories for UK thesis information, iv) Test the value of a national support service for e-theses creation and management in the UK, v) Work with other e-theses developments internationally, and in particular to assist the research aims of other e-theses projects funded within the JISC FAIR Programme, and vi) Produce a ‘checklist approach’ for universities to use as they develop e-theses capability.

The project was based around a core team of three staff at the University of Edinburgh- a Project Director, Project Officer and Systems Developer. The Project initially consulted with a group of pilot universities with the aim to take delivery of the *Theses Alive!* software for use in their institutions. As the project proceeded user feedback and product evaluation was gained from a larger community of institutions as part of the worldwide development and user group of our chosen software package.

Initially, the project carried out an extensive evaluation of the current open-source digital repository software available to the HE community. With a suitable underlying software platform chosen (DSpace) work then concentrated on building a bespoke digital repository and thesis submission system suitable for the requirements needed for the UK HE community. This software package, called TAPIR (*Theses Alive!* Plug-in for Institutional Repositories), is freely available to download as a self-contained add-on to the core DSpace code from the *Theses Alive!* web site, along with supporting installation documentation. The Tapir has now been downloaded and installed by several institutions, whose feedback has been instrumental in upgrading the software to newer versions. This development work from the *Theses Alive!* project, along with major input from the SHERPA project, has culminated in the creation of an Institutional Repository for the University of Edinburgh- the Edinburgh Research Archive (ERA).

Concurrently, a standard metadata schema for ETDs in the UK has been developed and subsequently has been implemented in the DSpace platform as part of the TAPIR software. This schema was developed in conjunction with the RGU *Electronic Theses* project, the GUL *Daedalus* project and representatives from the British Library.

During development of the software, and afterwards, the Project worked to provide a general information and user support service on ETDs. This service took the form of a mediated deposit service and ETD creation support. Practically this consisted of providing guidance for postgraduate students and supervisors on suitable file format types, scanning resolutions, conversion and system administration, through web-based technologies (email/web pages) or telephone support. This user support service was successfully piloted at the University of Edinburgh; however feedback from consultation with the pilot institutions indicated that a national user support service of this nature would not be appropriate. Instead we have successfully concentrated on disseminating our project findings primarily through the *Theses Alive!* website, published journal articles and conference papers.

In addition to these original aims and objectives set out in the project plan it has been necessary to investigate the effects on intellectual property rights (e.g. copyright and patents) and other legal implications (e.g. the Freedom of Information Act 2002) which arise when publishing research material online. These unforeseen problems proved to be a significant barrier to the progress of the project and the development of electronic theses programmes in general. The solutions delivered by the *Theses Alive!* project have been published by the JISC Legal service and have already proved to be extremely valuable to the HE community.

In conclusion, the e-theses service piloted by the *Theses Alive!* project has been warmly welcomed in Edinburgh. However it is clear that dedicated support from home institutions, in the form of changing the current thesis regulations to include provision for electronic submission, is required if e-theses programmes are to be successfully adopted.

Appendix B. Budget

<i>Two-year funding from JISC spread over 3 Financial Years</i>	YR1	YR2	YR3
	01/03 – 07/03	08/03 – 07/04	08/04 – 12/04
Income from JISC	16374.52	41502.75	26399.73
Add balance carried forward	16374.52	32964.29	14259.52
Staff			
John MacColl, 0.1 FTE @ AL5	2703.17	4819.00	2670.53
Theo Andrew, 0.5 FTE @ AL2	7024.50	13126.00	7273.99
Richard Jones, 0.75 FTE @ AD2	10536.31	19687.50	10910.16
Travel & Subsistence	4335.00	6746.00	1428.00
Equipment (items over £10k)	0.00	0.00	0.00
Dissemination activities	154.00	235.00	72.00
Evaluation activities	109.00	80.00	60.00
Other	51.00	411.00	0.00
Balance	-8538.46	-12140.21	-8155.16

Appendix C. Tapir Code Licence (BSD-style licence)

Copyright (c) 2003, The University of Edinburgh. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.*
- Neither the name of the University of Edinburgh, or the names of the contributors may be used to endorse or promote products derived from this software without specific prior written permission.*

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDERS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Appendix C. Theses Alive! website statistics: Document downloads

Rank	Item File Name	Description	Downloads
1	DspaceInstall.pdf	Technical Report	395
2	ETD2003report.pdf	Conference Report	268
3	Intellectual_Property_Rights_Workshop.pdf	Conference Report	262
4	SchemaComparison.pdf	Technical Report	301
5	ComparativeEvaluation.pdf	Technical Report	355
6	SherpaEdinburgh.ppt	Presentation	335
7	projectplan030305.pdf	Project documentation	255
8	projectplan030305.doc	Project documentation	290
9	MetadataSchemas.pdf	Technical Report	290
10	UKSG_Report.pdf	Conference Report	239
11	EdinburghResearchArchive.ppt	Presentation	331
12	Intellectual_Property_Rights_Workshop.doc	Conference Report	296
13	Trends.pdf	Technical Report	218
14	SparcEdinburgh.ppt	Presentation	306
15	DspaceEdinburgh.ppt	Presentation	301
16	DspaceInstall.doc	Technical Report	311
17	EThesesLeeds.ppt	Presentation	336
18	WorkflowOperation.pdf	Technical Report	277
19	OAF_thesesalive.ppt	Presentation	329
20	proposal030408.pdf	Project documentation	246
21	DspaceUGMeeting.ppt	Presentation	250
22	proposal030408.doc	Project documentation	274
23	ComparativeEvaluation.doc	Technical Report	279
24	ETD2003report.doc	Conference Report	278
25	SchemaComparison.doc	Technical Report	302
26	WorkflowOperation.doc	Technical Report	281
27	MetadataSchemas.doc	Technical Report	296
28	EThesesWorkflowProposal.pdf	Technical Report	210
29	OA_Forum_Workshop_4_9_03.pdf	Conference Report	223
30	FAIR_X4L_report.pdf	Conference Report	243
31	biannual030305.pdf	Project documentation	249
32	FAIR_X4L_report.doc	Conference Report	247
33	eFAIR_Minutes.doc	Project documentation	225
34	UKSG_Report.doc	Conference Report	234
35	biannual030305.doc	Project documentation	236
36	Trends.doc	Technical Report	225
37	eFAIR_Minutes.pdf	Project documentation	214
38	OA_Forum_Workshop_4_9_03.doc	Conference Report	226
39	ERAInstallation.pdf	Technical Report	189
40	ThesesSubmissionProcedure.pdf	Technical Report	205
	Total		10827

Table 2. Top 40 downloaded documents from the Theses Alive! site archive.

Appendix D. Tapir download statistics

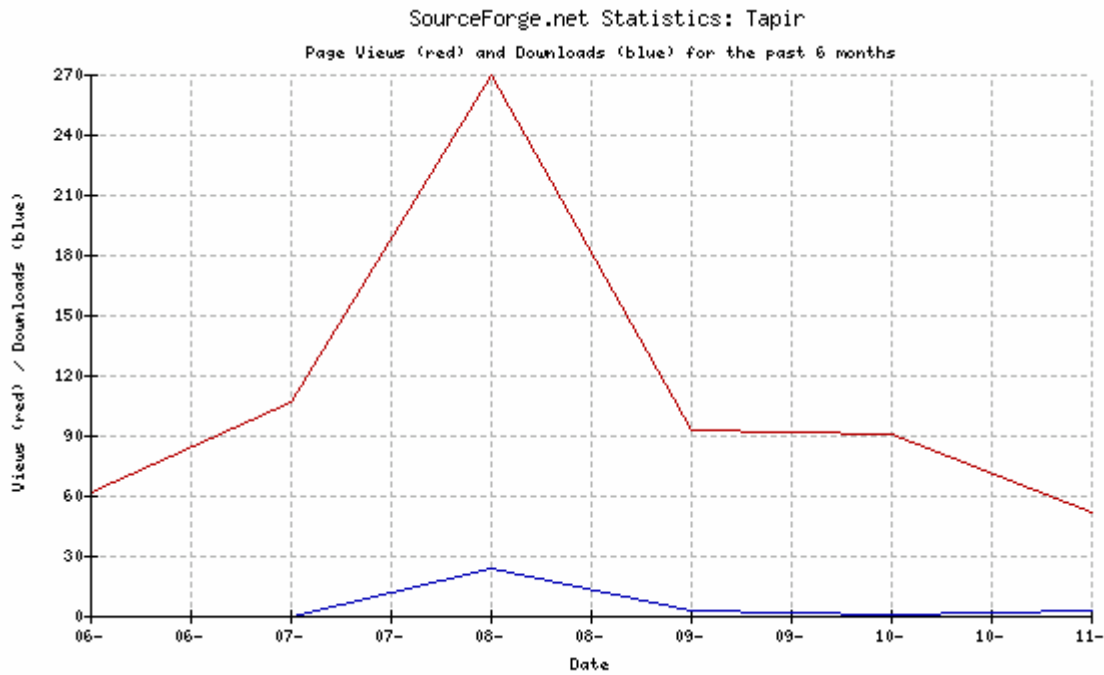


Fig. 1. Graph showing page views and downloads for the Tapir software (from <https://sourceforge.net/projects/tapir-eul/>)

Month	Page views	Downloads
Nov-04	52	3
Oct-04	91	1
Sep-04	93	3
Aug-04	270	24
Jul-04	107	0
Jun-04	62	0

Table. 3. Monthly breakdown of page views and downloads of the Tapir software (from <https://sourceforge.net/projects/tapir-eul/>)

Appendix E. ERA usage statistics: June-Sept 2004

General Overview

=====

Number of Bitstream Views:	35048
Number of Item Views:	23213
Number of Collection Views:	5337
Number of Community Views:	1894
Number of Logins:	132
Number of Searches Performed:	2363
Number of Theses Archived:	189
Number of Licence Rejections:	0
Number of OAI Requests:	99