JISC DIGITISATION PROGRAMME

Project Document Cover Sheet

UK Thesis Digitisation Project

PROJECT PLAN

Project

Project Acronym (if applicable)	UKTD
Project Title	UK Thesis Digitisation Project
Start Date	01/04/07	**End Date**	31/01/09
Lead Institution	British Library
Project Director	Andy Appleyard
Project Manager & contact details	Anthony Troman, British Library, Boston Spa, Wetherby, West Yorks, LS23 7BQ
Partner Institutions	CURL, Cranfield University, The University of Warwick, University of Glasgow, The University of Edinburgh, Robert Gordon University, The University of Birmingham
Project Web URL	Not available. EThOS project website: http://www.ethos.ac.uk/
Programme Name (and number)	JISC Capital Programme: DIGITISATION
Programme Manager	Alasdair Dunning

Document

Document Title	Project Plan		
Reporting Period	N/A		
Author(s) & project role	Anthony Troman, Project Manager		
Date	24/09/2007	**Filename**	UKTD Project Plan 20070924
URL			
Access	Project and JISC internal		

Document History

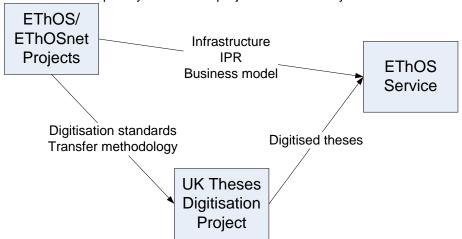| Version	Date	Comments
20070924	24/09/2007	First draft
20071009	09/10/2007	Improvements following JISC suggestion. Numbered paragraphs

JISC

JISC Project Plan

Overview of Project

1. Background

1.1 The project will digitise a minimum of 20,000 theses (specific number dependant on the cost to digitise per thesis).

1.2 Digitised theses will be made available via the EThOS service which is being developed by the EThOS and EThOSnet projects, projects jointly funded by JISC, CURL, The British Library and the development partner Higher Education Institutions. The EThOS service is expected to go live during 2008 – the actual live date is subject development progress and launch date considerations.

1.3 The relationships between the EThOS project, the EThOS system and this UK Thesis Digitisation Project are as follows:

1.3.1 The EThOS Project delivers a fully operational, easily scaleable and financially viable **prototype** UK e-theses online service that will enable end-users, via one single web interface, to access the full text of electronically stored theses after selection from a database of UK theses. The project also addresses the major issues surrounding the service such as the business model and IPR, and deliverables offer practical support to UK HE institutions in the establishment of Institutional Repositories and participation in the service.

1.3.2 The EThOSnet project will take the deliverables from the EThOS project and scale them up into full operation.

1.3.3 The EThOS service will implement the outcomes of the EThOS and EThOSnet projects as a live service.

1.3.4 The UK Theses Digitisation Project will digitise at least 5335 theses to the standards and methodologies developed by the EThOS project and will be loaded to the EThOS service for onward supply to Researchers. The theses loaded will be subject to the business model and IPR recommendations developed by the EThOS project and offered by the EThOS service.

1.4 The digitisation project is closely allied to the EThOS project and will increase the number of theses available in the early days of the service. Since the theses to be digitised are likely to be among the most 'popular', the load on those digitising in response to orders will be reduced and supply times improved. Further, the digitised theses from this project will 'seed' the EThOS service generating a 'critical mass' to encourage future submission of theses in electronic form. Digitised theses will be made available to those host HEIs which have their own Institutional Repositories should they wish to accept them. Without the theses digitised by the UK Thesis Digitisation Project,

the service will be less attractive to the Researcher (less theses available for immediate access i,e. to download) and supply times for theses digitised from paper will be greater since the digitising facility will be under heavier use.

1.5 Further information on the EThOS project can be found by following this link:

http://www.ethos.ac.uk/

1.6 The website for the EThOSnet project is currently under development. EThOSnet expects to deliver a live EThOS service for the start of the 2008 academic year. Further work will continue on the project until May 2009.

2. Aims and Objectives

2.1 The project will digitise at least 20,000 doctoral theses sourced from UK HEIs which will be selected based on British Library Thesis Unit and institutional usage figures, and other robust sources such as international experience e.g. NDLTD.

2.2 Selection criteria for theses to be digitised by this project are:
1) *Doctoral theses only*
   This is a constraint on the EThOS service and, since the digitised theses will be made available by EThOS, they must be consistent with the rest of the content offered.
2) *British Library Thesis Unit usage statistics*
   The British Library maintains a database of ALL theses supplied since the start of the service in the early 1980s and how many times each thesis has been supplied. This data can be used to generate a list of theses ordered by 'popularity'.
3) *Institution usage statistics*
   Institutions also maintain usage statistics on their theses. For example, the University of Nottingham has recently generated its list of 100 most 'popular' theses, and the University of Edinburgh has recently digitised a number of theses to make them available on their Institutional Repository. Where available, these statistics will be merged with the data from the British Library to generate a combined list of theses by usage.
4) *Availability of an e-copy*
   Only theses for which an e-version can NOT be obtained will be digitised. Any thesis already supplied to an Institutional Repository in electronic form (whether e-born or digitised) will be harvested by the EThOS service once live, so these theses will NOT be eligible for digitisation by this project.
5) *Age of thesis*
   Since the majority of theses supplied are from the last 13 years and these theses should also be robust and in good enough condition to withstand the handling required for digitisation, the figures generated in the above two criteria will be generated only for the last 13 years.
6) *Condition*
   Only theses which are robust enough to withstand the manual handling required for digitisation will be accepted.

2.3 Selection is based purely on 'popularity' i.e. the most requested theses within the criteria specificed above. The behind this is to reduce workload on the digitisation service at the British Library in the early days of the system by making the most popular theses available for immediate download.

3. Overall Approach

3.1 The project will use a commercial bureau identified as the result of a competitive tender exercise following EU rules to digitise the theses. The theses will be digitised to meet digitisation standards recommended by the EThOS Project and those used by the British Library for supply of other types of digitised material.

3.2 This type of competitive tender is mandatory under EU rules for the sums involved. The British Library is experienced at such tenders with its Corporate Procurement Unit (CPU) specialising in this area. The project manager is also experienced at working on this type of tender, subsequently delivering successful projects.

3.3 The advantage of such a tender to the project is that the 'competition' drives down prices and it is very likely that more theses will be digitised for the requested money. The disadvantage is that quality can suffer if a supplier cuts costs too far. The optimum supplier will offer a 'fit for purpose' quality service at a competitive price. The chosen supplier will be the one whose offering is judged the best value for money, not necessarily the cheapest.

3.4 The contract will be structured such that regular quantitative and qualitative checks will be undertaken to ensure that standards are maintained and that if problems arise the ultimate sanction of withdrawal from the contract and appointment of another supplier is an option.

3.5 Resource discovery, ordering and delivery will be via the EThOS service, which will offer further digitisation of paper theses on-demand by Researchers and the potential for further digitisation projects (see Appendix D – Business Model).

4. Project Outputs

4.1 The primary tangible deliverable from the project will be the digitised theses which will be loaded onto the EThOS system.

4.2 Technical and procedural knowledge and experience will be gained from working with the commercial digitising agency and will be applied to the on-going digitisation work undertaken at The British Library for the EThOS service.

5. Project Outcomes

5.1 The project will deliver at least 20,000 digitised theses. The competitive tender under EU rules is likely to force down the digitisation cost per thesis by putting potential suppliers into a competitive bid where digitisation standards and timescales are fixed leaving price as the only variable. This number is considerably more than originally estimated and has been achieved as a result of a competitive tender.

5.2 The contract with the supplier will be set up such that further digitisation work at the same price could be done should further funds become available e.g. an institution asks for further theses to be digitised at their own cost.

5.3 The digitised theses will be loaded to the EThOS service so offering some digitised content on day one.

5.4 The most 'popular' theses are being selected so that the initial load on the digitising service is reduced as much as possible. The most labour intensive part of the EThOS service is digitisation of ordered theses once they have been supplied to The British Library digitisation studio and this is therefore a potential 'bottleneck'. By digitising the most ordered theses and placing them on-line on day one, the likelihood of a bottleneck is reduced and supply times of theses are improved.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
HE teaching and research staff	• Digitised thesis • Greater visibility of their students' graduate research output • Guaranteed preservation of theses in perpetuity by the BL	High
HE postgraduates	• Digitised thesis • Greater visibility of their graduate research output • Guaranteed preservation of theses in perpetuity by the BL	High

Project Acronym:
Version:
Contact:
Date:

HE undergraduates & life-long learners	• Immediate access to the most popular 5335 theses on day one. • Guaranteed preservation of theses in perpetuity by the BL	Medium
Non-HE based and international users	• Immediate access to the most popular 5335 theses on day one. • Guaranteed preservation of theses in perpetuity by the BL	High
HE senior administrators	• 'Free' digitisation of selected theses • Theses supplied back to the institution (where they have Institutional Repositories. • Promotion of institutional research output • Enhanced visibility for their institution • Guaranteed preservation of theses in perpetuity by the BL	High
HE / BL information professionals	• Most popular theses available on day one • Improved resource discovery and document supply • Guaranteed preservation of theses in perpetuity by the BL • Space saving	High

7. Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
EThOS project does not deliver	2	5	10	Tight project structure and project management. Digitised theses could be stored on other BL systems until the project delivers or an alternative solution is found. In the case of total failure, the theses will be supplied to Institutional or Consortia Repositories.
Bureau does not deliver digitised theses.	2	3	6	Full and detailed examination of candidate digitisation bureaux as part of the EU tender process. Early and frequent examination of bureau work. Start project as soon as possible (including pre- award preparation) to allow delivery of theses as early as possible.
Digitisation is more expensive than quoted.	3	1	3	Supply bureaux with sample material. Use experience of the EThOS digitisation workpackage to ensure good and accurate analysis of bids. Fixed price contracts for fixed numbers of theses.
HEI refusal to participate.	1	1	1	EThOS partners listed on the cover sheet will participate. If other HEIs do not wish to take part, the project will identify the next theses on the list and digitise from that HEI. However, as the HEIs are receiving 'free' digitised theses it is expected that many will participate.
Staffing	1	1	1	The British Library has a large number of staff with the appropriate skills to meet the administrative requirements of the project.
Organisational	1	1	1	The project manager is very experienced at running complicated projects – comparatively, this is a simple project. Only digitisation bureaux with experience of similar material will be considered as partners.

Project Acronym:
Version:
Contact:
Date:

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Financial	1	1	1	The EThOS project delivered detailed metrics on the cost of digitising theses – this experience will allow good and accurate judgement of bureau costing proposals. Administrative tasks are completed at no cost to the project. Should digitisation costs exceed the expected price the number of theses digitised would be reduced – the project budget will not be exceeded.
Technical	1	1	1	The technical infrastructure for storage and delivery is being developed by the EThOS project (see above). The digitisation standards and data transfer methods are tried and tested from the digitisation partners on this project.
External suppliers	1	1	1	Evaluation criteria will include requirements that competing bureaux have extensive experience of either theses or similar types of material. Use of BL and project manager experience.
Legal	3	5	15	The EThOS and EThOSnet projects are addressing the legal aspects of collecting, digitising and making this type of material available.
More theses to be digitised than expected	5 (realised)	2	10	The project has delivered 4 times the number of theses to be digitised than originally expected. This means greater logistical involvement for the British Library, but the additional resource can be made available.
Institutions attempting to clear rights with authors	3 (small number realised)	3	9	A small number of institutions are contacting authors for clearance to make their theses available. 3 or 4 of the bigger institutions are doing this impacting on the logistics of the project. This is containable by applying a time limit of late May for decisions and the addition of further theses to replace those withdrawn..

8. Standards

8.0.1 Standards defined in this section are as expected at the start of the project. However, should competing bureau offer positive contribution by suggesting improved standards while delivering for the same cost, the proposals will be considered. Digitisation standards will be:
- Scanning at 300 dpi. This is the CAPTURE figure.
- Individual page images to be saved as .tiff
- Compression of tiffs with LZW (lossless compression) – possibly jpeg2000
- Colour pages to be held as 24 bit tiff images
- Grey scale pages to be held as 8 bit tiff images
- Black and white pages to be held as 1 bit tiff images
- Multi-page pdfs with OCR'd hidden text generated from the 24 bit, 8 bit and 1 bit tiff images

8.0.2 Final delivery format will therefore be the multi-page pdfs with OCR'd hidden text generated from the tiff images.

Document title: JISC Project Plan
Last updated: April 2007

Project Acronym:
Version:
Contact:
Date:

8.1 Data Transfer structure and naming convention

8.1.1 The naming convention for data transfer and subsequent storage on the EThOS system is well defined and was used successfully for the digitisation elements of the EThOS project.

8.2 Metadata

8.2.1 Other than file i.d.s and the specified file structure, no descriptive metadata will be supplied with the digitised items. All such metadata will be assigned on loading an item either by matching to an existing record generated by the EThOS project or by creating manually or automatically on loading.

8.2.2 Records from The British Library Integrated Library System will be imported to EThOS to initially form the bulk of the database. Records will then be downloaded from Institutional Repositories or continue to be catalogued in the current manner on the British Library ILS and subsequently imported to EThOS.

8.3 Preservation

8.3.1 Digitisation projects conducted at the BL must pass through the Digitisation Approvals Process which provides a set of guidelines and requirements to ensure digitisation projects are conducted to a sufficient standard.

8.3.2 The longevity of the BL's digital collections and digital assets and the maintenance of continued access to them is a critical part of the British Library's strategy. Key staff from the Digitisation Programme, the Digital Preservation Team and eIS/Operations are tasked with ensuring the longevity of the outputs of digitisation activities. This includes assessment, verification, the inventory of metadata, media checking and refreshment, and storage at multiple geographical locations.

8.3.3 With the recent establishment of a new digital preservation team and the ongoing development of a state of the art, distributed digital repository under the Digital Object Management Programme, the BL is aiming to provide an even higher degree of confidence in the provision of long term digital preservation. The digital masters from the UK Thesis Digitisation Project will ultimately be stored in this new digital repository which will ensure preservation by:

- Storing multiple copies of each digital object in different physical locations and utilising backups and integrity checking to ensure that no data is lost
- Characterising and validating each digital object in order to determine requirements for its preservation.
- Recording metadata about each digital object and allocating unique, persistent identifiers
- Developing and executing preservation plans
- Implementing a comprehensive technology watch mechanism to provide warning of the need to take preservation action.

8.4 Accessibilty

8.4.1 The EThOS service uses an interface originally designed and still in use for the search, order and delivery of e-journals and conference papers. The interface conforms to British Library standards for accessibility and design. All British Library web resources aim to meet W3C accessibility standards level AA and all British Library applications aim to meet best practice in usability of web resources. All major web developments are subject to lab based usability testing.

8.4.2 In terms of the theses themselves, supply of pdfs with hidden OCR'd text offers opportunities for those Researchers with visual impairment to use software to improve accessibility e.g. to increase the display size or use text reading software. The British Library has, on request, supplied digitised theses for those with visual impairment in exactly the format that these standards specify, with adequate results.

8.5 Compliance with the JISC Information Environment

8.5.1 The requirements specified by the JISC Information Environment are addressed by the EThOS project, which is offering access, storage and delivery functionality for the digitised theses. EThOS conforms with the requirements of the JISC Information Environment.

8.6 Compliance with MINERVA Digitisation Guidelines

The images delivered will meet the MINERVA digitisation guidelines.

9. Technical Development

9.1 The project does not involve technical development. The only technical standards involved are those defined for the digitisation (see section 8)

10. Intellectual Property Rights

10.1 Intellectual Property Rights for UK theses is complicated and led to lengthy discussions surrounding this project. The issues and approach are addressed in Appendix C. Theses digitised by the UK Thesis Digitisation Project will be subject to the IPR conclusions drawn by the EThOS project in full consultation with UK HE partners.

Project Resources

11. Project Partners

11.1 The partnership agreement with the selected commercial digitisation bureau will take the form of a formal contract to deliver a fixed number of theses for a fixed price to a specified timescale. The digitisation partner has not yet been selected so contact details cannot be given. Once selected, contact details will be available from The British Library project manager (Anthony Troman).

11.2 The HEI partners listed on the cover sheet are those from the EThOS project. Partner institutions will NOT contribute to funding of the project other than the staff salaries required to develop lists of most popular theses (if institutional statistics are available), retrieval of theses from the shelf, postage, packing and despatch to the British Library, and re-shelving theses post-digitisation i.e. exactly the same as all participating institutions. The partnership involved in this project primarily demonstrates support from the institution for the concept and the execution of this project which, in turn, helps the development of the EThOS service. The partner institutions have, or have in development, Institutional Repositories so would generally receive back their own digitised theses. Contact details are not listed as the partners are largely inactive in the project, but details are available from The British Library project manager.

11.3 In addition, there will be 'informal' partnerships with HEIs contributing theses to this project.

12. Project Management

12.0.1 The project will be managed within The British Library's Operations and Services Directorate under the Project Directorship of Andy Appleyard (Head of Operations and Customer Services).

12.0.2 Alasdair Dunning of JISC will be invited to join the board to represent the sponsoring body.

12.0.3 Due to the relatively simple processes involved in this project, reporting will primarily be in writing via email with Project Management meetings called should exceptions arise. Reports will focus on throughput rates of thesis digitisation and quality.

12.0.4 Key personnel forming the project team are as follows:

12.1 Anthony Troman – Project Manager

12.1.1 Anthony is a project manager qualified to PRINCE2 Practitioner level with 17 years practical experience of successfully delivering projects in a library environment and beyond. He has worked for 13 years at the British Library developing systems including full UKMARC bibliographic cataloguing, acquisition and access. He is workpackage leader on 3 of the 8 EThOS project workpackages (Central Hub development, Digitisation and Business Model) and successfully led the Digitisation workpackage. He sits on the EThOS Project Board and will take the delivered system into full operation at the British Library should all parties agree that the service should go live. Anthony has

successfully worked with contracted companies in developing systems at the British Library including procurement under EU rules.

12.2 Chris Dixon – British Library Contracts Manager

12.2.1 Chris Dixon has worked at the British Library for over 10 years with extensive procurement and contracts experience. Chris manages the Contracts Team which has a responsibility to undertake Library procurement in a professional and ethical manner operating under the Chartered Institute of Purchasing and Supply code of practice. In particular provides advice and guidance to internal customers on all aspects of EU procurement ensuring that the Library is fully compliant with current EU regulations. Chris is also a member of the Chartered Management Institute.

12.3 Carol Burton – British Library Admin. and Support Staff Manager

12.3.1 Carol Burton has worked at the British Library for 12 years, managing the British Library Thesis Unit for the last year. Within her team she has vast experience of the current British Library thesis service and her section maintains links with those responsible for thesis supply at the vast majority of UK HEIs. Carol will manage the collection of theses from HEIs, forwarding theses to the bureau and return of theses to the originating HEI. The British Library Thesis Unit under Carol successfully undertook this task for the EThOS project.

12.4 Paul Wheatley – British Library Digital Preservation Manager

12.4.1 Paul Wheatley is a specialist in digital preservation and works for the British Library. He has played a leading role in various national and international digital preservation activities, most recently the JISC funded "Life Cycle Information for E-Literature" project. Paul is currently the Digital Preservation Manager in the eIS Directorate of the British Library, and a member of the team responsible for ensuring the longevity of the UK's scientific and cultural digital heritage.

13. Programme Support

13.1 There are no specific areas where support from the programme manager will be required other than the general support required by a project manager working to JISC standards for the first time.

13.2 It must also be understood that this project is working within constraints which may not apply to other projects on the programme i.e. the service parameters defined by the EThOS and EThOSnet projects e.g. metadata standard, digitisation standard, IPR approach, business model, etc..

14. Budget

14.1 The budget is listed in Appendix A. The agreed project proposal will be adhered to. A figure of 5335 digitised theses was originally estimated. On awarding the project the bid was reduced, though the aspiration of having 5335 theses digitised remained. The reduction was made by JISC on the basis that a competitive tender would result in lower digitisation costs per thesis. As a result of the competitive tender, in fact 20,000 theses will be digitised – the considerable increase as a result of improved technologies and locating a considerably cheaper bureau than those experienced previously.

Detailed Project Planning

15. Workpackages

See Appendix B

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
WP1	Digitisation Partners	Experience	Request for	Against absolute

Project Acronym:
Version:
Contact:
Date:

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
	1	Qualifications Health & Safety Equality Etc.	information	criteria specified in RFI and scored against each other.
WP1	Digitisation Partners 2	Quality of work Cost	ITT	Best value for money (cost vs. quality)
WP1	Digitisation	Quality of project work	Quality process	See Quality section
WP2	Thesis selection	Most popular theses	Computer system report	No. of times supplied in last 13 years
WP3	Equipment	Most appropriate Size	BL experienced staff	Quality equipment holding all theses.
WP4	Workflow	Keep the agency working	Planning & constant monitoring	Digitised theses delivered to timescale

17. Quality Plan

17.1 The major deliverable from the project will be digitised theses and the quality plan will be based on this deliverable. The delivery schedule is yet to be addressed with the supplier (who is yet to be selected), but the following criteria will be applied in checking the theses.

- Sampling will be based on General Inspection Level II (Normal Inspection) as per the ISO-2859-1 standard (Sampling Procedure for Inspection by Attribute). Selection of files for sample will be randomly generated.
- Formal quality visits – 2 weeks after delivery of the first paper theses to the bureau there will be a formal visit to check procedures and quality of digitised theses, ensuring that theses are being produced to the expected quality, throughput rates are as expected and that early problems in any of these areas are 'nipped in the bud'. Half way through digitisation a further visit will take place.
- Number of theses digitised –procedures developed for the EThOS project will monitor the number of paper theses requested from an HEI, the batches in which each paper thesis was received, in which it was despatched to the bureau, in which it was returned from the bureau and the Lacie disc on which the digitised thesis was returned from the bureau. Digitised theses will be retained at the bureau and will be stored at the British Library for the lifetime of the project.
- Number of pages per thesis – programmes developed as part of the EThOS project will count numbers of digitised tif pages per thesis and will be used to test a random sample of every batch returned for tif page numbers against page numbers in the paper thesis.
- Pdf pages versus tif pages – a random sample of every batch will be tested to ensure that the pdf returned has the same number of pages as there are tif page files for the same thesis.
- Quality of scanning – each batch will be tested to ensure that the tifs and the pdfs have been scanned at the specified scan quality.
- Colour – a random sample of digitised theses in each batch will be compared with the original thesis for colour scanning of the appropriate thesis pages.

17.2 Although a limited amount of the quality checking can be done by computer (tif page counts, file sizes, etc.) much of the checking must be done by human eye. British Library Thesis Unit staff are very experienced at the quality checking of microfilmed theses and what to look for when quality assuring filmed theses. The scanning process for paper theses is similar to filming and the types of errors in some cases will also be similar so, although the viewing mechanism is different, British Library Thesis Unit staff can use their experience to identify potential errors.

18. Dissemination Plan

18.1 One of the major workpackages of the EThOSnet project is an advocacy workpackage. This workpackage is addressing all aspects of thesis service provision for the UK and is drawing up plans for the run up to the live date of the service at the time of writing. The Thesis Digitisation project is an important part of the service and will benefit from the advocacy work from this workpackage. There will be no separate advocacy plans for the individual development projects of the future thesis service, though if specific pieces are required for the Thesis Digitisation project are required they can be offered.

19. Exit and Sustainability Plans

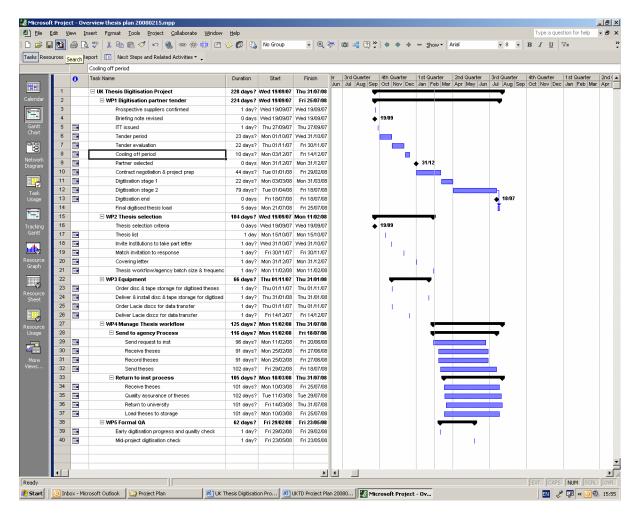| Project Outputs	Action for Take-up & Embedding	Action for Exit
20,000 digitised theses	Loading to the EThOS system for day one live operation	Maintained on and offered for delivery by the EThOS service and preserved by the BL Digital Object Management System.

Appendixes

Appendix A. Project Budget

The total grant awarded to the project by JISC is £451350.

Project Acronym:
Version:
Contact:
Date:

Appendix B. Workpackages

Microsoft Project - Overview thesis plan 20080215.mpp

File Edit View Insert Format Tools Project Collaborate Window Help

Cooling off period

#		Task Name	Duration	Start	Finish
1		⊟ **UK Thesis Digitisation Project**	**228 days?**	**Wed 19/09/07**	**Thu 31/07/08**
2		⊟ **WP1 Digitisation partner tender**	**224 days?**	**Wed 19/09/07**	**Fri 25/07/08**
3		Prospective suppliers confirmed	1 day?	Wed 19/09/07	Wed 19/09/07
4		Briefing note revised	0 days	Wed 19/09/07	Wed 19/09/07
5		ITT issued	1 day?	Thu 27/09/07	Thu 27/09/07
6		Tender period	23 days?	Mon 01/10/07	Wed 31/10/07
7		Tender evaluation	22 days?	Thu 01/11/07	Fri 30/11/07
8		Cooling off period	10 days?	Mon 03/12/07	Fri 14/12/07
9		Partner selected	0 days	Mon 31/12/07	Mon 31/12/07
10		Contract negotiation & project prep	44 days?	Tue 01/01/08	Fri 29/02/08
11		Digitisation stage 1	22 days?	Mon 03/03/08	Mon 31/03/08
12		Digitisation stage 2	79 days?	Tue 01/04/08	Fri 18/07/08
13		Digitisation end	0 days	Fri 18/07/08	Fri 18/07/08
14		Final digitised thesis load	5 days	Mon 21/07/08	Fri 25/07/08
15		⊟ **WP2 Thesis selection**	**104 days?**	**Wed 19/09/07**	**Mon 11/02/08**
16		Thesis selection criteria	0 days	Wed 19/09/07	Wed 19/09/07
17		Thesis list	1 day	Mon 15/10/07	Mon 15/10/07
18		Invite Institutions to take part letter	1 day?	Wed 31/10/07	Wed 31/10/07
19		Match invitation to response	1 day?	Fri 30/11/07	Fri 30/11/07
20		Covering letter	1 day?	Mon 31/12/07	Mon 31/12/07
21		Thesis workflow/agency batch size & frequency	1 day?	Mon 11/02/08	Mon 11/02/08
22		⊟ **WP3 Equipment**	**66 days?**	**Thu 01/11/07**	**Thu 31/01/08**
23		Order disc & tape storage for digitised theses	1 day?	Thu 01/11/07	Thu 01/11/07
24		Deliver & install disc & tape storage for digitised	1 day?	Thu 31/01/08	Thu 31/01/08
25		Order Lacie discs for data transfer	1 day?	Thu 01/11/07	Thu 01/11/07
26		Deliver Lacie discs for data transfer	1 day?	Fri 14/12/07	Fri 14/12/07
27		⊟ **WP4 Manage Thesis workflow**	**125 days?**	**Mon 11/02/08**	**Thu 31/07/08**
28		⊟ **Send to agency Process**	**116 days?**	**Mon 11/02/08**	**Fri 18/07/08**
29		Send request to inst	96 days?	Mon 11/02/08	Fri 20/06/08
30		Receive theses	91 days?	Mon 25/02/08	Fri 27/06/08
31		Record theses	91 days?	Mon 25/02/08	Fri 27/06/08
32		Send theses	102 days?	Fri 29/02/08	Fri 18/07/08
33		⊟ **Return to inst process**	**105 days?**	**Mon 10/03/08**	**Thu 31/07/08**
34		Receive theses	101 days?	Mon 10/03/08	Fri 25/07/08
35		Quality assurance of theses	102 days?	Tue 11/03/08	Tue 29/07/08
36		Return to university	101 days?	Fri 14/03/08	Thu 31/07/08
37		Load theses to storage	101 days?	Mon 10/03/08	Fri 25/07/08
38		⊟ **WP5 Formal QA**	**62 days?**	**Fri 29/02/08**	**Fri 23/05/08**
39		Early digitisation progress and quality check	1 day?	Fri 29/02/08	Fri 29/02/08
40		Mid-project digitisation check	1 day?	Fri 23/05/08	Fri 23/05/08

Ready

Document title: JISC Project Plan
Last updated: April 2007

Project Acronym: UKTD
Version: 1.0
Contact: Anthony Troman
Date: 24/09/2007

Workpackage and activity	Earliest start date	Latest completion date	Outputs (clearly indicate deliverables & reports in bold)	Milestone	Responsibility
WP1 *Digitisation partner tender*					
1. Prospective suppliers confirmed	Wed 19/09/07	Wed 19/09/07	Set of digitisation partners for ITT		AT/CD
2. Briefing note revised	Wed 19/09/07	Wed 19/09/07	Requirements statement for digitisation		AT/CD
3. ITT issued	Thu 27/09/07	Thu 27/09/07	ITT issued to prospective suppliers		AT/CD
4. Tender period	Mon 01/10/07	Wed 31/10/07	Deposited tender documents		AT/CD
5. Tender evaluation	Thu 01/11/07	Fri 30/11/07	Selected partner		AT/CD
6. Cooling off period	Mon 03/12/07	Fri 14/12/07	Consideration of selected partner		AT/CD
7. Partner selected	Mon 31/12/07	Mon 31/12/07	Confirmed selected partner	Y	AT/CD
8. Contract negotiation & project prep	Tue 01/01/08	Fri 29/02/08			
9. Digitisation stage 1	Mon 03/03/08	Mon 31/03/08	Set of digitised theses for 1st FY		AT
10. Digitisation stage 2	Tue 01/04/08	Fri 18/07/08	Set of digitised theses for 2nd FY		AT
11. Digitisation end	Fri 18/07/08	Fri 18/07/08	Completed set of digitised theses	Y	AT
12. Final digitised thesis load	Mon 21/07/08	Fri 25/07/08	Last thesis loaded.		AT
WP2 Thesis selection					
1. Thesis selection criteria	Wed 19/09/07	Wed 19/09/07	Specified BL system selection criteria		AT
2. Thesis list	Mon 15/10/07	Mon 15/10/07	List of 'most popular' theses		AT
3. Invite Institutions to take part letter	Wed 31/10/07	Wed 31/10/07	Letter sent		AT
4. Match invitation to response	Fri 30/11/07	Fri 30/11/07	Matched responses		CB
5. Covering letter	Mon 31/12/07	Mon 31/12/07	Letter to go with thesis list to inst.		AT
6. Thesis workflow/agency batch size & frequency	Mon 11/02/08	Mon 11/02/08	Agreed workflow with digitisation partner.		AT/CB

Document title: JISC Project Plan
Last updated: April 2007

Project Acronym:
Version:
Contact:
Date:

WP3 Equipment					
1. Order disc & tape storage for digitised theses	Thu 01/11/07	Thu 01/11/07	Order despatched		AT
2. Deliver & install disc & tape storage for digitised theses	Thu 31/01/08	Thu 31/01/08	Server hardware installed	Y	AT
3. Order Lacie discs for data transfer	Thu 01/11/07	Thu 01/11/07	Order despatched		AT
4. Deliver Lacie discs for data transfer	Fri 14/12/07	Fri 14/12/07	Discs ready for use	Y	AT
WP4 Manage Thesis workflow					
Send to agency Process					
1. Send request to inst	Mon 11/02/08	Fri 20/06/08	Letter to request theses		CB
2. Receive theses	Mon 25/02/08	Fri 27/06/08	Theses at BL		CB
3. Record theses	Mon 25/02/08	Fri 27/06/08	Theses listed		CB
4. Send theses	Fri 29/02/08	Fri 18/07/08	Theses despatched to agency	Y	CB
Return to inst process					
5. Receive theses	Mon 10/03/08	Fri 25/07/08	Theses returned to BL with digitised.		CB
6. Quality assurance of theses	Tue 11/03/08	Tue 29/07/08	Theses digitised match paper theses		CB
7. Return to university	Fri 14/03/08	Thu 31/07/08	Theses despatched to originating inst.		CB
8. Load theses to storage	Mon 10/03/08	Fri 25/07/08	Theses loaded to EThOS	Y	CB

Document title: JISC Project Plan
Last updated: April 2007

Appendix C – Intellectual Property Rights considerations

C.0.1 The EThOS project included a workpackage which examined the Copyright and IPR issues surrounding theses. The information in this appendix is taken from presentations and reports produced by the EThOS project. It is important to note that the workpackage was led by a UK HE representative and the reports were produced by the same person. Expert advice has been taken from British Library legal representatives and Charles Oppenheim has been asked to comment on the findings. The EThOS Project Board has signed off the IPR report and recommendations.

C.0.2 IPR issues arise because the aim is to offer an up-to-date service which meets the needs and expectations of 'Google generation' users i.e. rapid identification, location and supply of information, including immediate download. Into the future, should the service be taken up and offered by The British Library on behalf of UK HE, it is hoped that Institutions will amend thesis submission rules and documentation to allow authors to specify whether they are happy for their work to be distributed via the Institutional Repository and the EThOS service. Until ALL Institutions do this, the risks and issues addressed by this Appendix will apply.

C.0.3 In considering the risks and issues it should be remembered that at the moment there are circa 500,000 paper theses in existence WITHOUT permissions to allow digitisation and, in the vast majority of cases, no up-to-date contact details for the authors.

C.0.4 IPR risks primarily come from two sources:
1. Author IPR – the author of the thesis holds the IPR in most cases. *[Note that in some cases the institution claims the copyright – it is not clear whether this would be upheld if the author objected and challenged this assertion]*
2. 3rd party IPR – many theses contain material taken from previously copyrighted items.

C.1 Author IPR

C.1.1 UK Copyright law clearly states that to retrospectively scan a work without consent is a breach of copyright. Based on losses to the copyright owner, the further dissemination of these scanned works would certainly raise damages in any action. Ominously, there is also a provision in copyright law which can be used when assessing damages which looks to the flagrancy of the breach. The initial recommendation from the EThOS project was that the EThOS service should be an "Opt-in" service i.e. the author's permission should be sought before the thesis is digitised, stored and made available via the EThOS Central Hub.

C.1.2 However, the "*opt-in*" scenario would be extremely problematic for a future central hub service, in that:
1. The time it takes to seek out the author to obtain his/her permission would be prohibitive for the delivery times of the thesis to the customer. In the current service where this has to be done, delivery times can be up to 6 weeks.
2. Staff time and other operational costs make this very expensive and therefore untenable

C.1.3 In practice, then, the EThOS service will actually put in place an 'Opt-out' service where an author can specify that, if they find their thesis made available via the EThOS service, the thesis is removed immediately.

C.1.4 Charles Oppenheim was asked to comment on the level of risk that this approach might offer and suggested the following formula to calculate it:

$$R = A*B*C*D$$

Where **R** is the financial risk,
A = chance of infringement
B = chance of awareness
C = chance of action

Project Acronym:
Version:
Contact:
Date:

D = is the financial cost

C.1.5 Looking at each of these in turn,

A is 1 (i.e., 100%) – there can be no doubt that scanning the entire dissertation is infringement.
B will not be particularly high. Many authors of theses will not be aware of the EThOS service, or it won't occur to them to search it. Others may be pleased. Assume 0.1 (i.e., 10%)
C is likely to be a similar figure. The majority of thesis authors will probably be quite flattered that their work is of potential interest. Assume 0.1 as well.
D is potentially a high number. The basis of any legal action will be the money saved by EThOS in adopting this approach, which a judge would set at a reasonable royalty fee, for a thesis this would probably not be high, but let us assume £500. EThOS would also have to pay legal fees, which could be very high – say £50,000. Indeed, there is a strong argument that no case should ever go to Court, but that if sued, EThOS should immediately offer to make amends and should have a contingency fund set aside for these circumstances. Nonetheless, even here, some legal fees may have to be incurred. So let us say the typical legal costs would be £2,000. No calculation can be realistically made for loss of reputation, so let's ignore that factor.

C.1.6 Thus **R** for any given thesis **is** 1 x 0.1 x 0.1 x 2,500, or **£25.**

C.1.7 In the case of author IPR the following risk management strategies will be adopted by the EThOS service on the advice of the EThOS project:

Strategy	Description	Accept/Reject
Acceptance	The risk of an author taking action against the service are known and are considered low. The impact is also comparatively low.	Accept
Prevention	Adopt an opt-in service	Reject
Contingency	Implement a rapid 'notice and takedown' policy i.e. if an author objects to the thesis being made available remove the thesis	Accept
Transference	Take out an insurance policy	Accept – to be investigated further
Reduction	Engage opinion, publicity and marketing.	Accept

C.2 3rd Party IPR

C.2.1 Even if permissions were granted by the author to scan and/or make available a copy of the thesis online there may be further complications in the form of third-party copyright material such as images, journal articles, data or other licensed content being held in the thesis. This practice has traditionally been tolerated as the primary audience for the material has been the examination process with limited external supply. In addition, the existence of multimedia components in e-theses makes this issue more important in the coming years.

C.2.2 For theses submitted in future, the EThOS project will include information in its toolkit to advise authors on how to clear copyright for 3rd party material, but the issue remains for existing theses (and those supplied in the future without such clearance).

C.2.3 Once again, for retrospectively scanned material it is not practical to check every single thesis. Further, expert staff would be needed to do so and that staff would be expensive and would add considerable delay to the process. However the likelihood of inclusion of 3rd party material likely to generate action is thought to be relatively low.

C.2.4 It is felt that the 3rd Party risk likelihood may be higher than that of author IPR breach action, and the Impact also higher, but, given the nature of the material, not significantly so.

Project Acronym:
Version:
Contact:
Date:

The same risk management measures as those listed for Author IPR are appropriate for 3rd party IPR, with the main mitigation measure for both being that of an advocated, visible and advertised (on the "live" service site) "Fast Take Down" Policy.

C.2.5 The policy will be in context of introductory text that explains the intentions and "ethos" of EThOS i.e. to showcase individual's research; to market their institutions research; to market the UK's primary research; to make theses more accessible to those who need them; to therefore make knowledge more freely accessible throughout the research community, and to up and coming researchers; to contribute to the global knowledge pool.

C.2.6 However, the other mitigation measures such as insurance and a compensation fund may have a heavier financial weighting. This will be informed by an EThOS Risk Assessment analysis undertaken by Charles Oppenheim.

C.3 Conclusions

C.3.1 There are risks associated with supply of theses since the service is, strictly speaking, knowingly infringing copyright. However, there are the following mitigating factors:

- The service is cost-recovery i.e. no-one is financially benefiting from supply of the theses.
- The intention is to promote the author's and the institution's reputation rather than to profit.
- Procedures such as rapid take-down policies will be put in place to remove any thesis unintentionally offered against the wishes of the author or institution.
- Procedures such as rapid take-down policies will be put in place to remove any thesis offered which infringes 3rd party copyright on request.

C.3.2 Advice from British Library legal staff suggests the following in relation to 3rd party indemnity for those supplying or funding the service:

C.3.3 "The British Library, as a Crown Body, does not hold commercial insurance but self insures under the provision of the government indemnity scheme. HM Treasury rules do not permit The British Library as a public body to give a blanket indemnity **for any third party action**. In order to provide such an indemnity, The British Library, in common with other government bodies, would have to take out additional insurance only by demonstrating that we can achieve value for money by so doing. This would involve making a business case, based on risk to our indemnity, financial status, value of contract and other ratios, in other words - it becomes a convoluted and complicated process.

C.3.4 Bodies supplying theses or contributing to the funding of the service are likely to be governed by similar rules. Ultimately, whether third party indemnity can be offered rather depends on various freedoms of action under Funding Agreements with DCMS. At the very least, such indemnity is conditional on DCMS consent. However, it is hard to imagine how those supplying theses or funding the service could suffer loss as a result of acts or omissions of the British Library."

C.3.5 It is therefore not possible for The British Library as service providers to offer indemnity to suppliers or funding bodies, but it is unlikely that any action taken against the service could impact on such bodies.

C.3.6 It should be noted that neither the author of the bid nor any other team member on the EThOS project can think of an instance where an action has been taken against any thesis supplier anywhere in the world (including The British Library).

C.3.7 The British Library will consider these risks when deciding on whether to take up the EThOS service.