Project Acronym: OCIMCO
Version: Version 1
Contact: Gillian Evison
Date: 28 Sep 2009

JISC Project Plan

Overview of Project

1. Background

While UK universities are rich in collections of manuscripts relevant to Middle Eastern and Islamic Studies, access to and exploitation of these holdings is not particularly easy and has severely impeded UK researchers. It was in recognition of the high priority placed on electronic access to manuscript resources that Oxford's Centre for the Study of the Book brought together academics and librarians in a Curator's Round Table in 2007[1] in order to share news about developments in online access to Islamic manuscripts. The importance placed on easy electronic access to information about manuscripts by the UK academic community was further highlighted in the 2008 JISC funded review of user requirements for digitised resources by Exeter University, where digitisation of UK Islamic manuscript catalogues was also given a high priority by respondents[2].

The combined holdings of Oxford and Cambridge form the second largest collection of Islamic manuscripts in the UK after the British Library and are of considerable intellectual significance, including many early and rare items with a broad subject coverage including literature, religion, philosophy, poetry, mathematics, astronomy and medicine. Both libraries have been collecting Islamic manuscripts since the 17th century and still continue to acquire manuscripts by donation or through purchase, a notable recent acquisition for the Bodleian being the *Book of Curiosities,* a late twelfth- or early thirteenth-century Arabic manuscript containing a series of rare maps and astronomical diagrams (MS. Arab. c. 90)[3].

Much of the intellectual significance of the collections can be attributed to the fact that they have been built around donations and purchases made from members of the two universities. One of the most notable of these at Cambridge is the collection of Professor E.H. Palmer (1840-1882) who acquired many manuscripts, including Qur'an fragments, during his travels in Sinai and other parts of the Middle East. Cambridge University Library also holds the collections of Professor E.G. Browne (1862-1926), who collected manuscripts in his travels in Persia in the early 20th century and Professor Reynold Nicholson (1868-1945), who was a keen collector and added many fine examples to the library. Notable collections of manuscripts at Oxford include those of Edward Pococke (1604-91), Robert Huntington, (1637-1701) and Narcissus Marsh (1638-1713). Edward Pococke was the first holder of the Laudian Chair of Arabic at Oxford and his collection of over 400 volumes includes such treasures as the *Book of Roger* by al-Idrisi, dated 1553 (MS. Pococke 375), with its attractive world and regional maps, and the charmingly illustrated animal fables of Bidpai, copied in Syria in 1354 (MS. Pococke 400). One of the most significant items in the Huntington collection of over 600 volumes, is the illustrated twelfth-century manuscript on weaponry commissioned by Saladin for his own library (MS. Huntingdon 264) and amongst the 700 manuscripts that were bequeathed to the Bodleian by Marsh is the Arabic version of the Conics of the Greek geometer Apollonius of Perga (MS. Marsh 667). Dated 1070, this manuscript was used by the astronomer Edmund Halley for his 1710 edition of Apollonius's work.

[1] http://www.bodley.ox.ac.uk/csb/roundtable_islamicmss.htm
[2] http://www.jisc.ac.uk/media/documents/programmes/digitisation/islamic_studies_report_pdf.pdf , p. 20
[3] http://cosmos.bodley.ox.ac.uk/hms/home.php

Although Cambridge's catalogues are available online in a simple PDF format[4], they possess serious shortcomings. Access to information about Oxford's manuscripts is even more limited, being confined to print catalogues and card indexes. The published catalogues are not easy to navigate with numbering systems that are difficult to understand and indexes that are of limited value. Since catalogue publication dates range from 18th to the 19th century, transliteration conventions vary considerably and the Bodleian's *Bilbliothecae codicum manuscriptorum orientalium…pars prima…* by Joannes Uri (Oxford: Clarendon Press, 1787) is in Latin. Several hundreds of manuscripts came to Oxford and Cambridge after the publication of the printed catalogues and records of these are only available to scholars able to visit the libraries and search the card catalogues on site.

2. Aims and Objectives

This pilot project proposes to create c.10,000 summary records of manuscripts in the Oxford and Cambridge collections using an extensible TEI P5 schema based on that developed for manuscript description by the Enrich project[5] that will provide a crucial framework to which more detailed descriptions, transcriptions and digital images can be attached at a later date. Oxford will be able to make use of tools and in-house encoding expertise gained through previous projects, such as the JISC funded *Shakespeare Quartos Archive* in order to support Islamic curators at Oxford and Cambridge in the creation of these records. The records will be made available through a searchable interface on the Oxford University Library Services website but the project would also offer to contribute them to the European Manuscriptorum, where they would join the 10,000 manuscript records that are due to be mounted there by the National Library of Turkey. Close liaison with Islamic Manuscript Association (TIMA)[6] will be maintained throughout the life of the project. Contacts made at Oxford's Round Table of 2007 with projects such as the Princeton and University of California initiatives will enable the sharing of experience and good practice between institutions.

3. Overall Approach

3.1 Methodology

3.1.1. Content Creation

Oxford has a complete record of its Arabic manuscripts (including several hundred records for items not described in any of its published catalogues) in a card catalogue. In the first instance cards will be digitised to produce JPEG or TIFF files using Capita Total Document Solutions. This company has previously worked with Oxford on JISC funded John Johnson Collection Electronic Ephemera project and with many of the UK's other leading cultural heritage institutions. Oxford will then outsource metadata creation from these digitised cards to AMA DataSet Ltd. This supplier has worked with the Bodleian on a number of previous projects including the ArtSTOR project[7] and the Bodleian Incunabules catalogue. Cards will be re-keyed in UTF-8 and tagged using elements from the Enrich project's TEI P5 metadata schema Using such dedicated services provides both economies of scale and also clear guarantees about the quality and delivery of the images and metadata. The records will then be enhanced in-house by the Arabic curator, who will be seconded to this project, in order to update the transliteration to Library of Congress, provide Library of Congress subject headings and convert author names to forms used by the Library of Congress authority files.

3.1.2. Overall Technical Architecture

The technical architecture will be delivered by a systems developer at Oxford already on the staff establishment, who will be supervised by a Technical Manager and Technical Architect both of whose time will be allocated to the project as part of Oxford's institutional contribution. The project will make use of the Digital Asset Management System (DAMS) currently in use for digital library projects within

[4] http://www.lib.cam.ac.uk/deptserv/neareastern/catalogues.html
[5] http://enrich.manuscriptorium.com/
[6] http://www.islamicmanuscript.org
[7] http://www.bodley.ox.ac.uk/medievalimages/ & http://www.bodley.ox.ac.uk/csb/bod-inc.html

Oxford (notably for the Future Archive project and Oxford University Research Archive). This provides a robust and flexible architecture that can be readily adapted to changing demands and technologies over time as well as incorporating long-term archival and preservation capabilities. A key aspect of the architecture is that it permits, and expects, that there will be multiple applications which use and manipulate material within the DAMS – this project will develop several such applications.

3.2 Important Issues to be Addressed

- **Metadata Standards.** TEI P5 has been led by the Western Manuscript cataloguing community and the Enrich schema does not have the range of attribute values that would be needed to fully express the codicological features of Islamic manuscripts. While the basic level records will not be affected by this issue, it is important for the plans for future enhancements to ensure that the TEI framework will allow full expression of the codicology of Islamic manuscripts. The project team will work closely with other JISC Islamic projects to ensure that there is full agreement on any additional values and, with the help of JISC, feed these back into the TEI community to ensure future interoperability of Islamic manuscript cataloguing projects.

- **Script & Directionality.** The CSS for the Enrich project specifies a font that cannot support the Latin Extended Additional characters used by the Library of Congress Arabic transliteration standard and is left justified throughout, which is unsuitable for text that contains script in mixed directionalities. These issues will need to be addressed in order to produce readable output. The Digital Asset Management System at Oxford has not yet been tested for ingest of scripts with right to left directionality. There are clear synergies with the SOAS/Yale JISC/NEH project as Yale also uses FEDORA and this project will be able to benefit from the existing expertise at Yale gained through the OACIS and AMEEL projects.

- **Open Access.** All materials and metadata in the Open Access portion of the DAMS are fully accessible using OAI-PMH and OAI-ORE standards to maximise reuse in the wider community.

- **Evaluation.** Evaluation of the project will be built into its project management framework, consultation with JISC, quality review procedures, user-testing and analysis for publication and conference

- **Useability.** Usability testing will take place on the web interface and will involve a focus group of academics, students, librarians and curators.

3.3. Scope and boundaries

In order to keep the project within the limited time and budget specified by this call for proposals, Oxford will concentrate on creating metadata for its Arabic manuscript collection of some 5000 records. It is intended that the workflows developed for this project could be used to convert the 2780 entries for Oxford's Persian and Turkish texts at a later date.

This project will not be creating any digital images of manuscripts but the framework will allow links to existing scanned images where they exist. Cambridge University Library intends to add links to existing scanned images of its printed catalogue pages, so the additional information they contain can be accessed by users of the OCIMCO catalogue.

The methodology concentrates on creating a large number of minimum level records and enhancements, such as detailed codicological information, are outside the scope of the present project.

3.4 Critical Success Factors

- Timely agreement over metadata standards
- Timely modification of the tracking system to suit project requirements
- Timely development of a basic object model for Fedora and system administration for the project within the DAMS
- Retention of experienced curatorial staff
- Usability/accessibility and take up by the academic and student body

4. Project Outputs

OCIMOCO's agreed deliverables are:-

- Easy access to c.10,000 Islamic manuscript descriptions through a full-text faceted search engine, based on Apache-SOLR, that will allow searching in Roman and Arabic script

- An extensible catalogue that can accommodate further manuscript descriptions and allows for enhancement of descriptions to provide more detailed catalogue entries

- A web-based interface with features developed in consultation with the Academic user community giving access to predefined searches, such as Author, Recipient, Title, Library or other institution, and predefined browseable views based on these searches

- A website with a rich set of display features suitable for academic research, such as the ability to display alongside an item any annotations/footnotes/images and links to digitised versions, internal or external, where available.

- Agreed TEI P5 for the presentation of manuscripts online that can be extended to accommodate more detailed descriptions, transcriptions and digital images and shared with other Islamic Studie projects

- Cataloguing tools and cataloguing storage FEDORA-based catalogue storage solutions that can shared with other institutions

- Further interoperability through the availability of records via OAI-PMH exposure

- Enhanced visibility of collections through submission of records to the European Manuscriptorum

5. Project Outcomes

- Free online access to everyone to some 10,000 manuscript descriptions for two of the UK's most significant Islamic manuscript collections for use in teaching, research and outreach

- TEI P5 descriptive standards to share with the library and archive community that allow the development online Manuscript Catalogues with the rich level of description that would be found in a traditional printed catalogue raisonné

- A knowledge base of TEI P5 and digital assets within the Project Team shared freely with the wider Islamic Studies community

- Preservation of records currently only held in manual indexes as single copies without any electronic back-ups

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
Researchers with an interest in Islamic manuscripts or their content	Access searchable descriptive metadata.	High
Students and teachers at all levels who would use such material in teaching, learning and research if it were more easily available to them	Ability to browse without travelling to Oxford and Cambridge to search uncatalogued material.	High
Bodleian and Cambridge University Library	Improved discovery of collections, basis for producing enriched records and linking transcriptions and digitised content	High
The JISC and other funding bodies	Major interest in methodologies used to implement the project, and in ensuring that outcomes reflect the JISC's mission.	High
Advisory Group	Opportunity to shape the resource through regular review of project content and its potential impact on teaching, learning and research.	Medium

7. Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing				
Recruitment and/or retention of project staff	1	2	2	By opting to second subject specialists & technical developers already on the staff, recruitment and retention risks are low
Organisational				
Productivity of cataloguers	2	3	6	Since existing staff are involved in the cataloguing there is some potential to extend the cataloguing beyond the project if necessary.
Failure to meet long-term sustainability target	1	4	4	Both partners have demonstrated their commitment to the provision of digital resources in the long-term. Both Oxford and Cambridge have an on-going commitment to ensuring permanent access to its digital collections in perpetuity.
Technical				
Unsuitability of outputs	1	3	3	An evaluation plan that includes input from an Academic Advisory panel and user testing will ensure that the project meets the needs of the target user communities.
External suppliers				
Problems with metadata sub-contractor	1	3	6	The preferred choice of sub-contractor has successfully collaborated with the Bodleian on a number of projects. A contract will ensure that appropriate action can be taken if performance is unsatisfactory and a 20% contingency will be built into the delivery timeline.
Legal				
Breakdown of	1	5	5	The Project partners will adopt the JISC's

partnership				Consortium Agreement. Both partner organisations have a long-standing commitment to work on this project, and have worked successfully together in previous endeavours.
Intellectual Property Rights issues	1	4	4	IPR of metadata records will be managed through the consortium agreement.

8. Standards

Name of standard or specification	Version	Notes
Website Standards		
XHTML 1.0 (HTML 4.01)	http://www.w3.org/TR/xhtml1/	The website will be XHTML 1.0 compliant using HTML 4.01 features
Microsoft Word & PDF	Word 2003, PDF/A	The documents hosted on the OCIMCO website will be in Word and PDF formats (the latter for dissemination not archiving purposes)
OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)	http://www.openarchives.org/OAI/openarchivesprotocol.html	The metadata content of the catalogue will be available for harvest via OAI-PMH for incorporation into federated/harvest-based discovery tools.
OAI-ORE (Open Archives Initiative Object Reuse and Exchange)	http://www.openarchives.org/ore/	The content of the catalogue will be available for harvest via OAI-ORE to enable programmatic re-use. OAI-ORE can also be used for bulk transfer of content between Oxford and Cambridge.
Descriptive Metadata Standards		
TEI P5	Enrich http://tei.oucs.ox.ac.uk/ENRICH/index.xml	DTD may be modified to accommodate descriptive features of Islamic manuscripts not found in Western Manuscripts
Library of Congress Transliteration	http://www.loc.gov/catdir/cpso/roman.html	
Library of Congress Authority Files (Subject Headings & Name files)	http://authorities.loc.gov/	There will be instances where names of scribes or authors of manuscripts will not have entries in the Authority Files. Under these circumstances, names will be recorded using reference sources, such as Brockelmann's *Geschichte der arabischen literature*, and be formatted according to AACR2
Font: MS Arial Unicode		Unless an open-source font becomes available during the lifetime of the project that can display both Arabic and Latin Extended Additional characters, Cascading Style Sheets on the website

		will use Arial Unicode MS
RDF (W3C Resource Description Framework) and derivatives	http://www.w3.org/RDF/	RDF will be used to express the structure of the material within the catalog internally. The RDF for an item will also be extractable in derivative formats such as HTML, N3, N-triples and Turtle for interoperability with semantic Web applications.

9. Technical Development

All materials and metadata in the Open Access portion of the DAMS are fully accessible using OAI-PMH and OAI-ORE standards to maximise reuse in the wider community. Support for features such as RSS feeds, Zotero eCitation and integration with iGoogle are also provided as part of the basic feature set as a result of the Oxford University Research Archive development.

The project system will comprise the main components detailed below:

- **XML Editor Installation and Configuration.** Currently, oXygen is being investigated as a TEI entry tool and the project will later look to adopting the cataloguing tool that is being developed by the Islamic Manuscripts Project led by the Wellcome Library. The XML editor will either be implemented as an embedded part of the DAMS (and hence part of the online editorial interface) or, should this not prove to be possible, will be a standalone instance feeding into the data import.

- **Tracking Systems.** The project will require a system to track the process of digitisation and keying. In a previous digitisation project (the JISC-funded John Johnson Electronic Ephemera project), a web-based tracking system was developed using PostgreSQL and PHP which supports multiple users and sites. This can be adapted to suit the current project with relatively little effort.

- **Catalogue Storage.** This will make use of a portion of existing DAMS storage and object management capability and the objects stored will become subject to the Bodleian Libraries' general digital preservation processes as a result. Each catalogue item will be stored as a single Fedora (www.fedora-commons.org) object. The FEDORA object model is particularly flexible and well suited to this sort of project - which can potentially derive information from a diverse range of sources. Essentially, an object can be considered to be composed of a number of datastreams which are either representations of the object itself (images, TEI text, PDF's) or metadata describing the object. It does not proscribe particular file or metadata formats and can cope with objects with differing datastream compositions. Datastreams can also be added at any time. RDF is used to express relationships between objects and datastreams. As a result, items can be readily augmented with comments, additional data, attachments and external links without requiring architectural changes to the storage system. Material derived from different sources can be stored and indexed with their native metadata intact (as well as in a normalised form) so that no information is lost when importing catalogues from other sources. This is important for the long-term growth and viability of the resource. Layered over the FEDORA object store are a set of tools and services which provide full text indexing and faceted search (Apache-SOLR), XML query capability (EXIST), an RDF triple-store (Mulgara) along with administrative tools such as virus scanning, text extraction and job scheduling. Data sharing protocols such as OAI-PMH, OAI-ORE, Atom and RSS are also catered for. These services already exist as part of the DAMS architecture so the task of delivering the Catalogue comprises the following activities. A system administration is needed to set up appropriate access controls, service instances and accounts for the project. A basic object model for Fedora also needs to be devised that encapsulates the data from the project effectively and can accommodate the full Enrich schema to allow for creation of new records and enrichment of existing records, as soon as

either project partner is ready to move to the next level. This includes investigating appropriate metadata and RDF standards for describing objects and relationships; Implementing the object models, metadata standards and ontologies in FEDORA; Providing basic administrative/editorial interface.

- **Data Import.** The DAMS already has a system for harvesting/ingesting data from other sources (the "Combine Harvester" developed as part of the JISC BID project). This will be tailored to handle to ingest of the outputs from the cataloguing phase of the project. This will draw together elements from the XML editor, the tracking system and the catalogue

- **Website.** This will present the content of the system to users along with a set of tools to allow them to make best use of the material. The website will be delivered in collaboration with the project team as part of an iterative process that will include feedback from the Academic Advisory Group on look and feel and search features. It will offer a full-text faceted search engine, based on Apache-SOLR, enabling users to perform searches in Roman and Arabic script.

10. Intellectual Property Rights

Both project partners, as Libraries of Legal Deposit, have considerable experience in this area and a shared interest in demonstrating the highest possible regard for the IPR of other individuals and organisations.

Bodleian Library and Cambridge University Library will retain copyright of the metadata records created at each institution and the partnership agreement will give the Bodleian the right to store and deliver copies of the Cambridge metadata through the searchable interface. Both partners aim to provide the widest possible access to the metadata, free of charge as part of their service provision to the international scholarly community.

Both the Bodleian Library and Cambridge University Library are committed to the delivery and support of digital resources in the long-term, and it is in the interests of both to ensure that the outputs from this project are of the highest quality and are fully supported, thus offering the best prospects for their sustainability, reusability, and retention of value in the long-term. Both libraries are committed to ensuring that any digital resources resulting from its own collections must remain viable and accessible in perpetuity, and regard the on-going maintenance of data created as a result of this project as a necessary function of their service provision. Project outputs will be delivered online and free of charge by the Bodleian through its catalogue interface.

Project Resources

11. Project Partners

Role	Name	Main contact	Consortium Agreement
Project Partner	Cambridge University Library	Grant Young gy219@cam.ac.uk	Letter of Agreement Signed 28/8/09 (Appendix C)
Project Subcontractor (Oxford) for Digitisation of Catalogue cards	Capita Total Document Solutions Telford Road Bicester Oxon OX26 4UP	Tony Sapnik Tony.Sapnik@capita.co.uk T: 0845 604 0080 F: 01869 240 982	n/a
Project Subcontractor	AMA DataSet Ltd 50 School Lane,	Christine Beatty christine.beatty@ama.uk.com	n/a

(Oxford) for re-keying and TEI mark-up	Bamber Bridge Preston, Lancashire PR5 6QE	Tel: 01772 627534 Fax: 01772 627535	

12. Project Management

A Project Board, comprising representatives from Oxford University Library Services and Cambridge University Library will take responsibility for overseeing the project:

Board Member Role	Name & Job Title
Chair	Richard Ovenden Keeper of Special Collections (Associate Director), Bodleian Library, Oxford
Digitisation Services Representative	Michael Popham Head of the Oxford Digital Library, Oxford University Library Services
Digitisation Services Representative	Grant Young Digitisation and Digital Preservation Specialist, Cambridge University Library
Project Director& Project Manager	Gillian Evison, Head of Oriental Section, Bodleian Library
Islamic Subject Specialist	Catherine Ansorge/Yasmin Faghihi, Joint Heads of the Middle Eastern Department at Cambridge University Library.
Academic Representative	Dr. Judith Pfeiffer, Oxford University Lecturer in Arabic/Islamic History
Digital Archives Specialist	Susan Thomas, Digital Archivist, Bodleian Library, Oxford

Day-to-day management of the project will be shared between experienced senior staff at both partner organizations, with prime responsibility for liaison with the JISC resting with the Project Director in the first instance. There will be two project teams, one based at the Bodleian Library in Oxford and one based in Cambridge University Library.

Project Team

Project Role	Name
Project Manager & Project Director (Project Management 0.2 FTE)	Gillian Evison Head of Oriental Section Bodleian Library, Oxford gillian.evison@bodley.ox.ac.uk Tel. +44 (0)1865 277033
Cataloguing & Record Editing (Oxford)	Colin Wakefield Head of the Middle Eastern Collection Bodleian Library, Oxford Email: colin.wakefield@bodley.ox.ac.uk Tel. +44 (0)1865 277030
Cataloguing & Record Editing (Cambridge)	Catherine Ansorge Head of Near Eastern Department Manuscripts and Printed Collections Cambridge University Library West Road, Cambridge

	CB3 9DR UK Email: caa1@cam.ac.uk Tel.: +44 (0)1223 766368
Cataloguing & Record Editing (Cambridge)	Yasmin Faghihi Head of Middle Eastern Department Manuscripts and Printed Collections Cambridge University Library West Road Cambridge CB3 9DR UK Email: yf227@cam.ac.uk Tel.: +44 (0)1223 766368
Cataloguing & Record Editors (Cambridge)	Part-time assistants; to be appointed
Technical Manager (Oxford)	Neil Jefferies, OULS Systems & e-Resources Research & Development Project Co-ordinator (neil.jefferies@ouls.ox.ac.uk)
Metadata (Oxford)	Alexander Huber, Oxford Digital Library (alexander.huber@bodley.ox.ac.uk)

13. Programme Support

JISC is supporting a number of complimentary Islamic Manuscript initiatives and this project would benefit immensely from JISC facilitating joint meetings and workshops in order that teams develop solutions to common challenges. Where the projects do find new standards are necessary then JISC's influence within the wider community will help ensure that these are widely promulgated and reduce the risk of multiple, competing, standards emerging.

14. Budget – See Appendix A

Detailed Project Planning

15. Workpackages – See Appendix B

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Sep 09-Dec 09	Encoding guidelines	Can data from catalogues be matched with appropriate elements and attributes in the TEI P5 schema? Is the schema robust & extensible? Is it compatible with encoding in use by other JISC Islamic manuscript projects?	Review by Oxford University Computing Services *Enrich* Project Team Peer review by other JISC Islamic manuscript projects & TIMA	Consistency of record structure between project partners and with other JISC Islamic projects

			Cataloguing Subcommittee	
Sep 09	Quality & quantity of TIFF images	Are images of sufficient quality and are files organised logically?	Review by sub-contractor quality control/ Project Team	Meet required standard; delivered on schedule
Sep 09-Jan 11	Quality & quantity of metadata records	Delivery/timeliness. Fit for purpose meeting user need (curators, academics & students)	Review by Project Team/Board	Meet required standard; delivered on schedule
Apr 10-Jun 10	Integration of Tools with DAMS	Fit for purpose meeting needs of cataloguers & record editors	Review by Project Team/Board	Module meets user requirements
Sep 10-Jan 11	Website usability	That catalogue functionality that is relevant to target audience (i.e. HE and FE)	Stakeholder focus groups, 'soft launch' to gather early feedback	Take up by target audience, increased use of manuscript collections
Ongoing	Documentation		Peer review	

17. Quality Plan

Timing	Output	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
Sep 09	Scans	Fitness for purpose	Individual project profile, sampling during scanning to ensure images meet required benchmarks. Quality control team for additional checks and a Project Supervisor who re-checks at least 10% of all images	Images mounted on project website. Acceptance by re-keying and mark-up subcontractor	Quality control team at Capita & Oxford Project Team	Project profile, scanning benchmarks
Oct 09-Dec 10	Metadata	Accuracy of input	Double-keying of data and manual checking of output data	The two keying results will generate a log file	Quality control team at AMA & Oxford Project Team	Compare tool
Oct 09-Jan 11	Metadata	Adherence to metadata standards	Compliance with TEI P5, Library of Congress	Well-formed XML documents compliant	Project cataloguers & Project Metadata	Configured XML editor & LC authority

			Transliteration & LC Authority Files	with the project metadata schema	Specialist	files online
Nov 10	Website	Adherence to standards	Assess relevant parts of the website with standards test harnesses.	Logs from test suites	Oxford Project Team	W3C XHTML 1.0 W3C, University of Vermont Repository Explorer (OAI-PMH)…

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
Sept 09	JISC blog	All interested parties; JISC	Raise Awareness	Funding Secured
Oct 09	TIMA codicology workshop	Islamic manuscript specialists	Raise Awareness	Funding secured
Oct 09	Oxford Institutional newsletter	Local Institution	Raise Awareness	Funding secured
Oct 09	Project website launched	All interested parties	Raise awareness, tool for Project Team communication, keeping all interested parties up to date on developments	Updates on project news
Oct 09	JISC Project Website	All interested parties	Raise awareness and contribute to a community of interest in Islamic Studies Catalogue and Manuscript Digitisation programme	Raise awareness
Nov 09-Feb 11	Progress reports at Faculty meetings and Committees on Library Provision	Oxford & Cambridge Faculty Members	Raise awareness, create interest in usability testing	Raise awareness, recruit for usability testing
Nov 09-Feb 11	Progress reports and Presentations to Library Stakeholder Groups e.g. NACIRA & MELCOM	Librarians and other interested parties	Raise awareness	Share metadata standards and project outputs with curators of Islamic Manuscript Collections
Summer 10	Report at TIMA annual conference cataloguing	TIMA members	Raise awareness	Share metadata standards and

	special interest group, contributions to TIMA cataloguing Wiki			project news with TIMA members
Feb 2011	Second Curatorial Round Table event hosted by Oxford Centre for the Study of the Book on Islamic manuscripts online focusing on project outcomes	Members of the academic community, librarians and other interested parties	Raise awareness and promotion of project catalogue	Share project output, standards and methodology and raise awareness of value to users
Ongoing	Promotion via Oxford & Cambridge Libraries	Members of the academic community, librarians and other interested parties	Raise awareness; embed use in academic community	Raise awareness of value to users

19. Exit and Sustainability Plans

At Oxford, a member of the Library's Systems & e-Research Service (SERS) team will archive all the digital outputs of the project and implement suitable preservation procedures, in keeping with current standards and best-practices, to ensure the longevity of, and long-term access to, the data.

The project will make use of the Digital Asset Management System (DAMS) currently in use for digital library projects within Oxford (notably for the Future Archive project and Oxford University Research Archive). This provides a robust and flexible architecture that can be readily adapted to changing demands and technologies over time as well as incorporating long-term archival and preservation capabilities. All materials and metadata in the Open Access portion of the DAMS are fully accessible using OAI-PMH and OAI-ORE standards to maximise reuse in the wider community. Support for features such as RSS feeds, Zotero eCitation and integration with iGoogle are also provided as part of the basic feature set as a result of the Oxford University Research Archive development. Cambridge University Library will additionally maintain copies of its records within its institutional repository (DSpace@Cambridge).

Project Outputs	Action for Take-up & Embedding	Action for Exit
Repository of Islamic Manuscript descriptions with searchable web-front end	Promote catalogue through methods outlined in dissemination plan; infoskills training for reader service staff, academics and students via library and live demonstrations at promotional events	Ensure user instruction and technical documentation is written and up to date and available on website along with internal team documentation
Metadata	Ensure metadata conforms to standards and is relevant to target audiences (i.e. HE and FE) and enables effective searching of the catalogue	Ensure metadata support documentation is complete and accessible, via the web and within teams. Apply appropriate Creative Commons Licence
Tracking System	The tracking system is an open-source application based on PostGRESQL and PHP. The core components of the system are likely to be re-used for other digitisation projects (this project represents a re-	In the long term, this is a project-oriented function so that obsolescence is not an issue.

	use case already).	

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Catalogue Content	The content represents a key library digital asset and will therefore be managed and preserved as part of the digital collections.	The Oxford DAMS and DSpace@Cambridge are designed for the long-term preservation of digital materials. Over time it is expected that the collections will be extended and enhanced with further material and scholarship.	Ensure that future work builds on existing resources rather than starting anew, thereby introducing interoperability issues.
Web-front end to catalogue & repository	Essential tool for accessing the catalogue descriptions	Operational responsibility for web front-end for joint catalogue will remain with Oxford's SERS. Cambridge will additionally ingest its records within the DSpace@Cambridge institutional repository	Need to ensure that all technical documentation is up to date & training delivered to Cambridge Systems team
Metadata tools	Essential tool for creation of online Islamic manuscript descriptions and transferable to other manuscript collections using non-roman scripts (e.g. Hebrew)	Metadata standard will be shared with and promoted within TIMA, the UK libraries and archives community and in the TEI community	Need to ensure that documentation is complete and accessible
Documentation & Evaluation	Will be completed within the project timeframe and made available publicly via the project OCIMCO website. The findings of the project – its successes and the problems it faced – will be shared within the communities of interest to help develop best practice	Ensure documentation is up to date and available via website	None

Appendixes

Appendix A. Project Budget

JISC Project Plan Budget : OCIMCO

Directly Incurred [redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted] [redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Appendix B. Workpackages

JISC WORK PACKAGE

Before completing this template please note:

- *The Project Management Guidelines have detailed instructions for preparing project plans and work packages..*
- *Please expand tables as appropriate.*
- *Fill in the information for the header, e.g. project acronym, version, and date.*
- *Text in italics is explanatory and should be deleted in completed documents.*

WORKPACKAGES	Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1: Project Management		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
2: Encoding Guidelines		X	X	X	X																				
3: Integration of Tools & Digitization Processes				X	X	X	X																		
4: Outsourced digitisation, re-keying & mark-up		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									
5: In-house data entry						X	X	X	X	X	X	X	X	X	X	X	X								
6: Integration of Tools with DAMS									X	X	X														
7: Quality Assurance		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
8: Development of a web presentation of the content													X	X	X	X	X								
9: Evaluation & Dissemination					X	X	X	X	X	X	X	X	X	X	X	X	X	X							

Project start date: September 1 2009

Project completion date: February 28 2011

Duration: 18 months

				Milestone	Responsibility
WORKPACKAGE 1: **<u>Objective</u>: To ensure effective internal communications & success of the project**					
1. Project management activities	Sept 09	Feb 11	Manage progress of OCIMCO		GE
2. Create & approve JISC Project Plan	1 Sept 09	30 Sept 09	**Complete JISC Project Plan**	**M1**	GE
3. JISC website project page	1 Sept 09	30 Sept 09	Publication of **OCIMCO webpage** on JISC website		GE
4. Create OCIMCO website at Oxford	1 Sept 09	1 Nov 09	Publish **OCIMCO website** on OULS site		GE
5. Set up Project Team & Partner Meetings	Oct 09	Jan 11	Series of monthly project team meetings/ conference calls		GE
6. Set up Project Board meetings	Oct 09	Jan 11	Series of quarterly Project Board Meetings		GE
7. Monitor finance	Sept 09	Feb 11	Keep financial expenditure within budget		GE
8. Project closure		Feb 11	Closure of OCIMCO project		GE

WORKPACKAGE 2: ***<u>Objective</u>: Development of encoding guidelines***					
8. Selection of metadata model & customising as required	Sep 09	Dec 09	**TEI P5 metadata schema** suitable for use with Islamic Manuscripts	**M2**	AH/NJ
9. Install a suitable editor and configure for use	Sep 09	Dec 09	XML editor configured with project metadata schema		
WORKPACKAGE 3: **<u>Objective</u>: Integration of Tools & Digitisation**					
10. Customisation of JISC-funded John Johnson Electronic Ephemera project tracking system	Nov 09	Feb10	Project tracking system		NJ
11. System administration to set up appropriate access controls, service instances and accounts for the project	Nov 09	Feb 10	Accounts & access controls for project		NJ
12. Devising a basic object model for FEDORA that encapsulates the data from the project effectively and can accommodate the full TEI P5 schema	Nov 09	Feb 10	Fedora object model in place		NJ
13. Implementing the object models, metadata standards and ontologies in Fedora	Nov 09	Feb 10	Object models & metadata standards in place in DAMS		NJ
14. Provide basic administrative/editorial interface for catalogue	Nov 09	Feb 10	Basic structure of electronic catalogue in place	**M3**	NJ
WORKPACKAGE 4: Outsourced digitisation, re-keying and mark-up **<u>Objective</u>: Manuscript catalogue records**					

14. Digitisation of Oxford catalogue cards	Sep 09	Sep 09	**TIFF files of c.5000 catalogue cards**		GE
15. TIFFS prepared for re-keying and mark-up pilots	Oct 09	Feb 10	Mark-up workflow in place		CW/GE
16. Re-keying and TEI mark-up	Feb 10	Dec 10	Basic TEI records for c. 5000 catalogue card entries		CW/GE
WORKPACKAGE 5: In-house cataloguing **Objective:**					
16. Enhancement of re-keyed records	Feb 10	Jan 11	**TEI records for c. 5000 Islamic texts at Oxford**	**M4**	CW
17. Creation of TEI XML records from entries in printed and manual catalogues at Cambridge	Jan 10	Jan 11	**TEI records for c. 5000 Islamic texts at Cambridge**	**M4**	
WORKPACKAGE 6: **Objective: Integration of Tools with DAMS**					
18. The "Combine Harvester" developed as part of the JISC BID project will be tailored to handle ingests of the outputs from the cataloguing phase, drawing together elements from the XML editor, the tracking system and the catalogue	April 10	June 10	Archival copies of digital output of the project, in keeping with current standards		NJ
WORKPACKAGE 7: Quality assurance **Objective: To ensure that records, repository and website are fit for purpose**					
19. QA of catalogue card scans	Sept 09	Sep 09	TIFF files of suitable quality for re-keying		CW/ Project subcontractor
20. QA of re-keying & TEI mark-up	Oct 09	Dec 10	TEI P5 files compliant with metadata schema		CW/AH/ Project subcontractor

21. QA of catalogue records	Oct 09	Feb 11	Accurate catalogue records		CW/YF/CA
22. Usability testing of website catalogue	Sep 10	Jan 11	Catalogue that answers stakeholder needs		NJ/CW/WF/CA
WORKPACKAGE 8: **Objective: Development of Web Presentation of Content**					
23. Development of a full-faceted search engine, based on Apache-SOLR, enabling users to perform searches in Roman and Arabic script	Sep 10	Jan 11	Searchable **OCIMCO website** presenting content of the system to users along with a set of tools to allow them to make best use of the material.	**M5**	NJ
WORKPACKAGE 9: Evaluation & Dissemination **Objective: To review the success of the project and to widely disseminate outputs and findings**					
24. Publicise project	Dec 09	Jan 11	Presentations to stakeholders		GE/CW/YF/CA
25. User evaluation for the web-content and search functionality	Sep 10	Jan 11	Focus group of stakeholders		GE/CW/YF/CA
23. Publicise website and online catalogue	Jan 11	Feb 11	Presentation materials, live and demonstrations of resource to stakeholders		GE/CW/YF/CA
24. Progress Report 1 to JISC	Feb 10	Feb10	Submit **Progress Report 1** to JISC		GE
25. Progress Report 2 to JISC	Sep 10	Sep 10	Submit **Progress Report 2** to JISC		GE
26. Draft & complete Final Report to JISC	Feb 11	Feb 11	Submit **Final Report** to JISC	**M6**	GE

Members of Project Team:

GE – Gillian Evison – Project Manager
AH – Alexander Huber – Metadata Developer
NJ - Neil Jefferies - Technical Manager, OULS Systems & e-Resources Research & Development Project Co-ordinator
CW – Colin Wakefield – Cataloguing and record editing (Oxford)
YF- Yasmin Faghihi - Cataloguing and record editing (Cambridge)
CA- Catherine Ansorge - Cataloguing and record editing (Cambridge)

Appendix C. Letter of Agreement

RESEARCH SERVICES
University Offices, Wellington Square, Oxford OX1 2JD

Grant Young
Digitisation and Digital Preservation Specialist
Cambridge University Library
West Road
Cambridge
CB3 9DR

Letter of Agreement
JISC Grant Ref: IRCONISL: Oxford Ref: R14442/CN001
'Oxford & Cambridge Online Islamic Manuscript Catalogue'
Dr Gillian Evison

The University of Oxford ("Oxford"), whose legal title is 'The Chancellor, Masters and Scholars of the University of Oxford' and whose administrative office is at University Offices, Wellington Square, Oxford, OX1 2JD and The University of Cambridge, acting through the Cambridge University Library ("Cambridge"), whose legal title is 'The Chancellor, Masters and Scholars of the University of Cambridge' whose registered address is at The Old Schools, Trinity Lane, Cambridge CB2 1TN have agreed to collaborate on an award from the JISC ("the Sponsor").

The project is due to commence on 1st September 2009 for a period of 18 months. Cambridge will use all reasonable endeavours to carry out the work which the principal collaborators agreed in the original proposal should be assigned to Cambridge. Both parties will use all reasonable endeavours to abide by the terms and conditions of the Sponsor's award letter dated 27th July 2009, which is attached at Annex A.

Oxford will forward to Cambridge the total sum of **sixty seven thousand, seven hundred and twenty two pounds (£67,722)** towards the cost of the research, subject always to receipt by Oxford of the funds from the Sponsor. Payment will be made in accordance with the award letter and the breakdown of costs attached in Annex B. Reimbursement will be made against accurate quarterly cost statements quoting '**F5RQTK0**' on the invoice. The cost statement should include the full fEC figures as well as the actual sums claimed.

If the terms of this Letter of Agreement are acceptable, I should be grateful if you would arrange for the enclosed copy to be signed by an authorised signatory on behalf of Cambridge and returned to me.

Yours sincerely,

Linda Andrews
Research Services Manager

Signed on the behalf of Cambridge

Name: [illegible] Jarvis

Title: University Librarian

Date: 25/8/09