JISC

Project Document Cover Sheet

Project Information			
Project Acronym	MIRAGE		
Project Title	Middlesex medical Image Repository with a CBIR ArchivinG Environment		
Start Date	April 1, 2009	**End Date**	September 30, 2010
Lead Institution	Middlesex University		
Project Director	Dr. Xiaohong Gao		
Project Manager & contact details	Alex Chapman Centre for Learning and Quality Enhancement Tel: 0208 411 5339 Email: a.chapman@mdx.ac.uk		
Partner Institutions	None		
Project Web URL	http://image.mdx.ac.uk/mirage		
Programme Name (and number)	*JISC Information Environment: Repositories: start-up 02/09*		
Programme Manager	Amber Thomas		

Document Name			
Document Title	*Project Plan*		
Reporting Period	*for progress reports only*		
Author(s) & project role	Xiaohong Gao (project director), Alex Chapman (project Manager), and Janet Rix (repository evaluator)		
Date	28 April 2009	**Filename**	Mirage_ProjectPlan.pdf
URL	*if document is posted on project web site*		
Access	☐ Project and JISC internal	V	General dissemination

Document History		
Version	**Date**	**Comments**

Project Acronym: MIRAGE
Version: 1.0
Contact: Alex Chapman
Date: 27/04/2009

JISC Project Plan

Overview of Project

1. Background

This start-up repository project will develop a warehouse of medical images and facilitate effective online retrieval tools in the institutional web site to complement the existing online e-leaning and teaching system OASISplus, also known as Blackboard *Vista*, that is currently in operation at Middlesex University (MU). Although OASIS+, which is managed by our Centre for Learning and Quality Enhancement (CLQE), can upload images, it is the way in which images are indexed and thereby retrieved effectively that poses a challenge, which will be the focus in this project. The chosen platform is GNU GIFT, an open source that has been adapted in the community of archiving medical images. Embedding to and integration with OASIS+ is seen as the path to the sustainability of the repository. To benefit other institutions, the developed new interface and algorithms will be released under open-source licences and will be designed to be compatible with and to contribute to the JISC e-Framework, with which, full engagement is envisaged.

MU has recently introduced a new MSc programme in BioMedical Modelling and Informatics (BMI), a collaborative venture between two schools, Engineering and Information Sciences (EIS) and Health and Social Sciences (HSS), which is now attracting an increasing number of students. During the course of studying and conducting final projects, a large amount of images have been employed in addition to other forms of data. At present, those images are communicated via portable media (e.g, USB sticks) between tutors and students, which has caused tremendous inconvenience, given the volume size of the data (>20,000 images). Further concerns include the time consumed in selecting the right sets of data and potential security risk ascribed to memory sticks (e.g, worms). Delay is inevitable if the delivery is missed or the wrong set of data is copied. Although a small portion of images has been uploaded to OASIS+ for downloading, retrieving images is usually very different from text retrieval, the way to search OASIS, calling for a repository of medical images serving its own need.

2. Aims and Objectives

The overall aims of the project are to set up a multimedia repository of medical images and to facilitate a multi-modal retrieval tool, which will be met by the following objectives:

- Extract image contents following the MPEG-7 standard by using two descriptors. One is a contour-based shape descriptor (to describe lesions) and the other is a global feature of texture (to present a domain), which will be saved as XML document files to be compatible with the other web page standards;
- Annotate images with their diagnostic keywords, e.g. tumour, head injury, etc., based on the accompanying text of an image using the Unified Medical Language System (UMLS);
- Locate anatomical locations of diagnostic lesions in each image based on standard domain atlases, e.g. Talairach atlas for brain images;
- Build a lookup table to link keywords with their content features in terms of shape, texture, and anatomical location;
- Design a multi-functional interface to facilitate text-based, content-based and semantic-based image retrieval;
- Review issues on ethical, privacy and copyright in dealing with medical data;
- Promote an e-learning environment;

- Interface with OASIS+ based on SCORM.

3. Overall Approach

The project will be structured into 2 stages representing two kinds of activities the repository will offer. Stage 1 will deposit images into the repository with meaningful (semantic) tags, which allows multi-modal retrieval to take place in Stage 2. To achieve this, the project will be employing agile open source development strategies such as SCRUM, delivering weekly cycles of development. The project will be built on the existing image server interfaced by GIFT, an open-source GNU image-finding tool that will be expended to incorporate both CBIR and text-based retrieval techniques.

The main focus of the development will be the ingestion of the repository and the design of the infrastructure of an interface between MIRAGE and OASIS. The research teams in both EIS and CLQE will supervise the design and evaluation of the repository and will be working together with a web page developer to be employed in the project for the development of the repository. Work on the extraction of image contents from the collections, development of the framework of lookup table for semantic retrieval, and annotation of images using keywords and anatomy information based on the vocabularies of UMLS will be largely undertaken by two current PhD students as part of their research work. All the features extracted here will be saved as XML document files.

The existing image server (http://image.mdx.ac.uk) archiving PET brain images will be built on, and its access interface GIFT, an open-source GNU image-finding tool, will be expanded to incorporate both CBIR and text-based retrieval techniques.

4. Project Outputs

In the immediate term, MIRAGE will benefit the University MSc BMI programme in providing students with a rich supply of medical images to research and will deliver:

- A project plan;
- A repository with ~20,000 processed medical images of multiple modalities;
- A framework of semantic retrieval of medical images;
- A potential interface with OASIS+;
- Project web site;
- Comprehensive usability evaluations and research results on the vocabulary used for description of medical images;
- Dissemination activities.

In the long term, the project will supply:

- A collection of benchmark images for evaluation of existing and emerging approaches for feature extraction;
- A collection of students' research results;
- An open source software for semantic image retrieval; and
- Integration with text-based retrieval of e-learning systems, OASISplus.

5. Project Outcomes

- Promotion of sharing medical images for the purposes of research and education;
- A platform for researchers to evaluate their image processing methodologies.
- A test-bed for evaluation of existing algorithms for medical image processing;

- A platform for sharing medical images and research results to a wider community, especially in UK higher education institutions where ~80% are employing Blackboard Vista (aka OASISplus at MU) as their online learning support system;

6. Stakeholder Analysis

Stakeholder	Interest/stake	Importance
EIS	The established repository will provided improved access and comprehensive content that will be engaged with their MSc BMI programme, arriving at more successful teaching and learning outcomes. The research results gained from the project will enhance the research activities in the School and will attract wider publicity and higher number of MSc student intake.	High
CLQE	The online OASISplus e-learning system will be complemented by adding to its capacity and services without a corresponding increase in resource, leading to a higher quality of learning based on e-resources.	High
CIE	The developed repository will also serve the other campuses of MU, including those in the other countries, thereby accommodating equal accessibility.	High
JISC	The experience in archiving medical images could be re-applied in similar projects, and the software (as well as the interface with OASIS) could be easily shared in the UK HE/FE for education and research purposes since ~80% UK universities are adopting Blackboard Vista.	Medium

7. Risk Analysis

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk
Staffing – Risk of being unable to employ suitable staff in a suitable timescale	2	4	8	Special care has been taken in the implementation of WP 5 to be carried out by a web developer, which starts four months into the project, allowing sufficient time for recruitment. By then the two PhD students will have started WPs 2-4, forming part of their on-going PhD studies, paving the way for WP 5.
Organisational – The project requires collaboration and co-operation between the partners to ensure a successful outcome	3	3	9	The project will be managed according to the JISC project management guidelines. Three collaboration teams, i.e., EIS, CLQE and CIE are all located on the same campus of Middlesex University in Hendon, enabling close co-operation of the team so as to ensure the delivery of the project outcomes.
Technical – technical difficulties that prevent implementation and deployment of the technologies that have been selected	2	3	6	The project will integrate existing technologies where the collaborators already have experience within the field. CLQE has the expertise in OASIS+, whilst EIS has in-depth knowledge in CBIR and GIFT, and CIE will obtain users' requirement. The implementation will follow JISC standards closely and open source infrastructure to enable global accessibility, and is supported by skilful technical support staff.
Legal – inability to resolve the legal issues effecting the data to be archived	1	5	5	The medical image data collected so far are satisfied with informed consent. For example, part of collection (~10,000) come from University Hospital of Geneva and have been utilised as benchmark images for many years for Cross Language Image Retrieval (ImageCLEF). This minimises the probability of this risk being materialised.
Legal – inability to define or reach agreement on the	1	5	5	The project is structured to ensure that the institutional deployments do not cross-institutional boundaries minimising potential issues arising from

legal issues regarding the repositories.				cross organisational legal issues regarding content. Since policies are already in place for the use of the institutional repository within the institution this risk is minimised but not entirely eliminated.
Short timescale, slippage of work package schedule	2	3	6	Progress will be kept under constant review through regular team meetings, particularly in the early stages. Extra resource will be made available where slippage is identified. The project has strong institutional support.

8. Standards

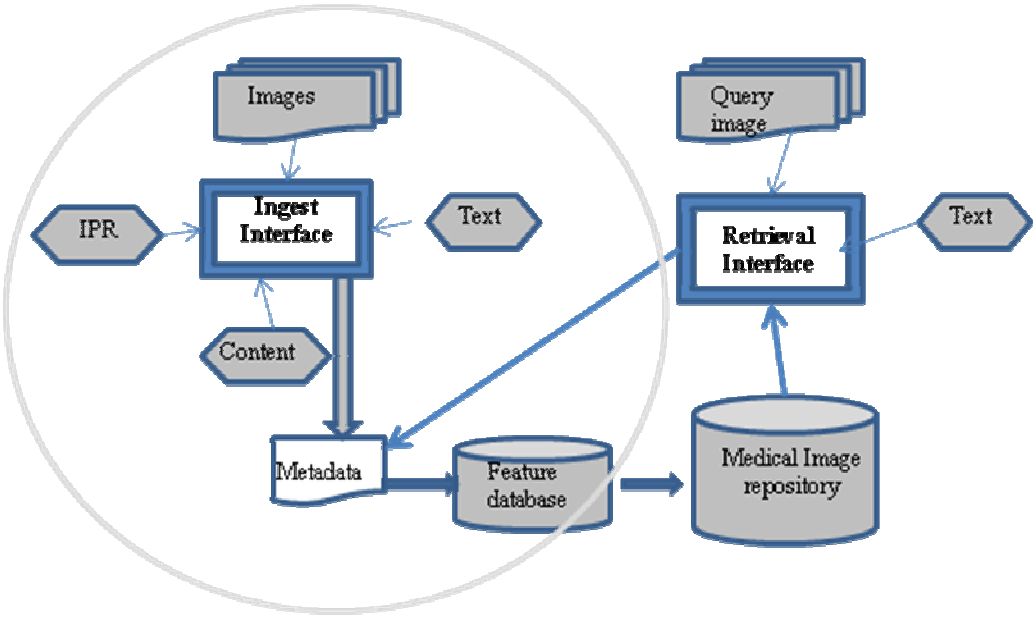
Name of standard or specification	Version	Notes
GNU GIFT	(2001)	http://www.gnu.org/software/gift/
Talairach brain atlas	1.1	http://www.talairach.org/applet/
MPEG-7	10	http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm
XML		http://www.w3.org/XML/
RedHat Linux System	7.1	http://www.redhat.com/
UMLS	5.0	http://www.nlm.nih.gov/research/umls/
SCORM	2004	http://www.scorm.com/

Quality Control on this project will include the use of quality reviews where the software, models and documentation will be reviewed by a subset of the team. Other elements of quality control will include the regular maintenance of a risk register and issues log.

9. Technical Development

The project will adopt the existing and emerging standards wherever possible and employ state of the art techniques in the fields of both content-based and text-based image retrieval. The ULMS will act as a gold standard to annotate images, whilst the MPEG-7 standard will be applied to extract the content of images. Application of standard atlases for locating ROIs will be practiced throughout the project as well as the adaptation of SCORM when interfacing with OASIS+ to be generic and compatible with Blackboard and other open source. The GIFT interface with a MySQL database will be employed together with OASIS techniques for the image server.

The follow diagram illustrates the repository to be developed. The main deliverables are highlighted on the left hand side within the circle whilst the rest of them are in place already.



10. Intellectual Property Rights

The software developed in the project will be released as open source under the GNU license policy to serve the wider community. All data contained within and associated with the images will be entirely anonymised and are satisfied with the consent form of medical data.

The copyright of any materials developed as part of the project will comply with the JISC requirements.

Project Resources

11. Project Staff

Xiaohong Gao (EIS) (Principal Investigator), mailto:x.gao@mdx.ac.uk
Alex Chapman (CLQE) (project manager), working on OASIS+, mailto:a.chapman@mdx.ac.uk
Martin Loomes (EIS), software development, mailto:m.loomes@mdx.ac.uk
Richard Comley (EIS), medical imaging, mailto:r.comley@mdx.ac.uk
Balbir Barn(EIS), software engineering, mailto:b.barn@mdx.ac.uk
Alex Moon (CLQE), OASIS+ and e-learning, mailto:a.moon@mdx.ac.uk
Janet Rix (CIE), evaluation, mailto:j.rix@mdx.ac.uk
Web page developer (to be recruited), repository development.

12. Project Management

The project will be collaboration between three parties at MU, including EIS, CLQE, and the Centre of International Education (CIE) to provide needed expertise. Regular meetings will take place to monitor the progress and day-to-day project management will be coordinated by the director.

A steering group will set up at the beginning of the project to monitor progress and to ensure ethical standards are observed when dealing with medical data. The IT support team in the School will be responsible for updating the server with the latest operating systems and to monitor the hard disk usage consumed by the growing volume of image data.

Internal and external communications will be managed by a dedicated (external facing) website with an integrated Blog and Wiki to support internal requirements. This structure is based on considerable JISC project experience for example Remora project.

The project will be executed following project plan tightly exactly.

13. Programme Support

The project is seeking advice and support from JISC project managers, training and dissemination opportunities through JISC, and potential future funding opportunities.

14. Budget

The budget plan is given at Appendix A.

Detailed Project Planning

15. Workpackages

The project has been structured into nine work packages that are detailed in Appendix B.

16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Sep. 2009	Physical state of image data being processed ready for depositing into repository.	If the collections are suitable to be processed? Any accompanying textual descriptions available for annotation?	Review by steering committee.	Selections of images between 10,000 to 20,000 will be processed.
Mar 2010	Usefulness of the repository	Is the repository helpful in searching needed information?	User group; peer group; analysis of usage logs of websites	Quality criteria defined by the steering group.
July 2010	User friendly interface	Is the content searchable, discoverable, and accessible in good time?	User group, project team and invited expert to access the interface.	Employment of appropriate standards; presentation of retrieved data meets users requirement.

17. Quality Plan

Output	Project Documentation – Plans and Reports				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
As of WP1 and WPs 8-9 (See	Compliance with JISC guidelines and	In accordance with published JISC guidelines	Consistent with JISC guidelines	Action by project team, review by steering committee.	

Appendix B)	expectations				

Output	Assessment of the contents of images being processed, annotated, and deposited.				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
May 09 - Dec 09 (WP2-WP5 as shown in Appendix B)	Comparison with well known good practice	In accordance with published JISC guidelines	Consistent with recognised quality standard, such as, TASI	Action by project team, review by steering committee.	
Ongoing (WP2 – WP5 as shown in Appendix B)	Feature extraction of images are fit on purpose	Review by academic experts	Project outputs consistent with project requirements	Project team.	

Output	Project web site; integration with OASIS				
Timing	**Quality criteria**	**QA method(s)**	**Evidence of compliance**	**Quality responsibilities**	**Quality tools (if applicable)**
Apr 09 – Sep 10 (WP6, WP7 as shown in Appendix B)	Consistent with relevant standards	Comparison with relevant standards	Meet with these standards	Action by project team, review by steering committee.	

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
Month 1	Project website; wiki; blogs;	Education community	Raise awareness and implications of the project on e-learning	Aim and progress of project; and key information relating to the project
Month 1	Public JISCMAIL list and other relevant lists	UK HE community	Share information and experience	Progress of the project
Month 6	Dissemination to conference of European Digital Library	various	Raise awareness and seek feedback on technical aspects	Project experience
Months 7-14	Completion of coding of repository	various	Share resources	Progress of project
Months 15-18	User evaluation	User group	Enhance repository	Raise awareness of repository
Months 1-18	Project reports, JISC meetings and workshops	UK HE community and related projects	Share experience with the other JISC funded start-	Project progress and experience

			up projects	
Month 18	Internal workshop	JISC; Middlesex University, users of OASIS+ and medical images	Publicise the project output, attract wider usage of the repository, and seek feedback.	Evaluate the repository.
Months 1-18	Publications to conferences and journals	JISC; wider national and international educational community	Publicise project findings and seek wider feedback.	Enhance the repository further according to the feedbacks.

19. Exit and Sustainability Plans

Exit plan

Project Outputs	Action for Take-up & Embedding	Action for Exit
Papers	publish	
Other documents	online	
Software	Online as open source	Accessible within MU and known institutions under mutual agreement on sharing medical data.
Mirage 1.0	MU BMI programme will use the repository for teaching and research, which will be embedded into or interfaced with MU e-learning system OASIS+.	Release Mirage 1.0 source code with full document to JISC start-up programme user community. The repository will be maintained by MSc students on BMI programme as part of their final project.

Sustainability plan

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
MIRAGE 1.0 – multi-modal interface for retrieving images	The algorithms used for content extraction are based on standard methods, whilst the interface is based on open sources being widely adopted in the field.	Staff/students of MU continue to contribute to the growth of the repository. Interfacing with MU OASIS is underway.	Interfacing with OASIS+ and accessing to Grid environment.

Appendix A. Project Budget

Directly Incurred Staff	Apr09–Mar10	Apr10 – Mar11	Apr11 – Mar12	TOTAL £
[redacted]				

Appendix B. Workpackages

WP1: Literature Review. This work package will review the state of the art in the related field and will inform the selection of technologies and their fit with the development of WPs 2,3,4,5. [D1][D3]

WP2: Image annotation: Each image will be described using keywords and anatomical information taken from the accompanying text with reference to vocabularies from the Unified Medical Language System (UMLS). Whereas the images without tags, usually used for control purposes, e.g., a normal brain image, will be left with blank text. [D2]

WP3: Image content extraction: This task is concerned with enhancing our currently developed CBIR system following the MPEG-7 standard by using two descriptors. One is a contour-based shape descriptor and the other is a global feature of texture to describe both local (ROI, region of interest) and global properties of an image. As for the anatomical position of a ROI, standard digital atlases will be applied. For example, in the domain of brain, the Talairach brain atlas will be applied, which is free online and can upload user images. [D2]

WP4: Image lookup table: To provide a certain level of semantic retrieval, a lookup table will be built to link text and contents together aiming at interpretation between the two modals (textual and content) of descriptions of an image and will employ a well-established Bayesian network.[D2]

WP5: Image ingest and retrieval interfaces: Built on GIFT, the interfaces will facilitate a metadata retrieval taking either or both text and a sample image as a query to provide a certain degree of semantic retrieval. Five distinct sources of information in computing the similarity distances between a query and images in the repository will be enabled: [D2]

- text-based objective term similarity (exact match)
- text-based content term similarity (inexact match)
- image-based objective term similarity (exact match)
- image similarity (inexact match)
- metadata-based (with combined textural and content) similarity (inexact match)

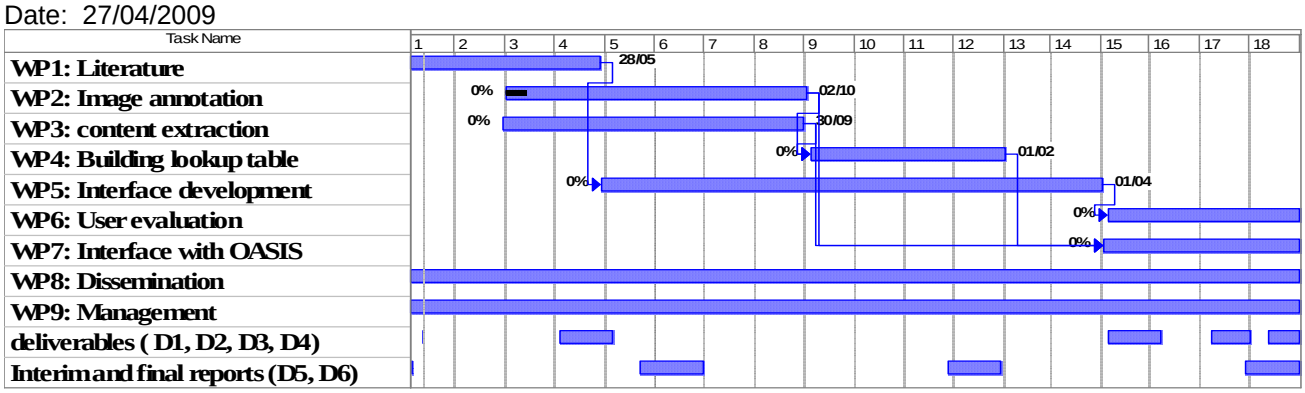
WP6: User evaluation: The interface design will adopt an iterative design methodology, which will deliver an improved understanding of the different classes of users that make use of medical images and their information seeking and retrieval behaviour. These users are primarily MSc and PhD research students at MU who use the repository as well as teaching staff. It is foreseen that wider community will make full use of the repository once it is in place. [D5]

WP7: Interface with OASIS. This package will be concerned with the potential integration with OASIS. Teamed with OASIS team at CLQE, an infrastructure will be designed based on SCORM, a sharable content object reference model. [D4]

WP8: Dissemination. This ongoing activity will disseminate the results of the project through appropriate channels to the wider community, a dedicated project web site, and publication and presentation at related workshops. [D6]

WP9: Management of the project will be the responsibility of EIS with significant input from CLQE and CIE into WP6, WP7, and WP8.

Project Plan


Task Name
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
WP1: Literature
WP2: Image annotation
WP3: content extraction
WP4: Building lookup table
WP5: Interface development
WP6: User evaluation
WP7: Interface with OASIS
WP8: Dissemination
WP9: Management
deliverables (D1, D2, D3, D4)
Interim and final reports (D5, D6)
28/05
0%
02/10
0%
30/09
0%
01/02
0%
01/04
0%
0%