JISC

| Project Information | | | |
|---|---|---|---|
| **Project Acronym** | CLIF | | |
| **Project Title** | CLIF (Content Lifecycle Integration Framework) | | |
| **Start Date** | 01 April 2009 | **End Date** | 31 December 2010 |
| **Lead Institution** | University of Hull | | |
| **Project Director** | Chris Awre | | |
| **Project Manager & contact details** | Richard Green<br>*r.green@hull.ac.uk* | | |
| **Partner Institutions** | Centre for e-Research (CeRch), King's College London | | |
| **Project Web URL** | www.hull.ac.uk/clif | | |
| **Programme Name (and number)** | IE Programme: Repositories Enhancement | | |
| **Programme Manager** | Amber Thomas & Andy McGregor | | |

| Document Name | | | |
|---|---|---|---|
| **Document Title** | Project Plan | | |
| **Author(s) & project role** | Richard Green (Project Manager) | | |
| **Date** | 24 April 2009 | **Filename** | CLIF-projectPlan-v10.pdf |
| **URL** | http://www.hull.ac.uk/clif/downloads/CLIF-projectPlan-v10.pdf | | |
| **Access** | ☐ Project and JISC internal | | ☑ General dissemination |

| Document History | | |
|---|---|---|
| **Version** | **Date** | **Comments** |
| v0.1 | 090409 | First draft |
| V0.3 | 090420 | Modified to include extra JISC requirements |
| V1.0 | 090527 | Additional full budget |
| | | |

THE UNIVERSITY OF HULL

and

KING'S College LONDON

# CLIF Project

## Project Plan

April 2009

## The CLIF Project

**Project Director:** Chris Awre (c.awre@hull.ac.uk)
**Project Manager:** Richard Green (r.green@hull.ac.uk)
**Project Site Manager for King's College:** Mark Hedges (mark.hedges@kcl.ac.uk)

The CLIF Project is being undertaken by the Information Systems Group at the University of Hull and the Centre for e-Research (CeRch) at King's College London. It is funded by the JISC Information Environment Programme 'Repositories Enhancement' strand.

# Overview of Project

## 1. Background

"No man is an island, entire of itself"
*John Donne (1572-1631)*
*Devotions Upon Emergent Occasions, Meditation XVII*

At the heart of meeting institutional needs for managing digital content is the need to understand the different activities that the content goes through, from planning and creation through to disposal or preservation. Digital content is created using a variety of authoring tools. Once created the content is often stored somewhere different, made accessible in possibly more than one way, altered as required, and then moved for deletion or preservation at an appropriate point. Different systems can be involved at different points: one of these may be a repository. To embed repositories in the content lifecycle, and prevent them becoming yet another content silo within the institution, they thus need to be integrated with other systems that support other parts of this lifecycle. In this way the content can be moved between systems as required, minimising the constraints of any one system.

The concept of a content lifecycle is not a new one.[1] Records managers have long recognised its importance to their work, and the JISC Supporting Institutional Records Management programme in 2003[2] looked to build on this and the previous Study of the Records Lifecycle project.[3] The MoReq2 specification[4] also refers to the document lifecycle in the context of electronic records management systems. In the commercial world of enterprise content management there is much consideration of how the content lifecycle can be improved to maximise the benefit the content offers a business, and there is no shortage of commercial offerings to enable this.[5] One of the issues that arises in this field, though, is the lack of standards for managing the content across systems.[6] Standards do exist (see section 3.2), though, and can be applied to this space. The Digital Curation Centre Curation Lifecycle Model[7] is being used to stimulate work on the use of standards to support this, and builds on an earlier piece of research at UKOLN.[8] This stated:

"The life cycle approach is necessary because:

- Digital materials are fragile and susceptible to change from technological advances throughout their life cycle, i.e. from creation onwards;
- Activities (or lack of) at each stage in the life cycle directly influence our ability to manage and preserve digital materials in subsequent stages;
- Reliable re-use of digital materials is only possible if materials are curated in such a way that their authenticity and integrity are retained."

Similar reasoning has seen the University of Illinois Library recently propose a digital content management approach.[9] The LIFE project also considered the content lifecycle, though explicitly left out hardware and systems from their cost models.[10] In contrast, the Information Technology and Information Storage industries (SNIA association) have addressed this issue, referring to the "… most appropriate and cost effective IT infrastructure …" required.[11]

The CLIF project will build on this previous work and address how to facilitate content lifecycle management within an institution.

## 2. Aims and Objectives

The CLIF Project will meet the four challenges of the JISC's funding call as follows:

2.1 By linking the repository into other content creation and management environments it will be taken upstream in the user's workflow. Where the repository is best positioned within the content lifecycle requires investigation: it may be relevant at the end of the creation stage to move the content into a repository for access and/or preservation; or it may be appropriate to move content into the repository as a staging area for subsequent processing.

2.2 The aim of integrating the repository at the appropriate part of the content lifecycle is to ensure that, when user activity crosses system boundaries, users do not feel constrained in what they wish or need to do; rather, the systems in question between them support these wishes and needs. For example, moving content used for teaching in a VLE into a repository, maybe as part of building a portfolio, supports content re-use and the potential for long-term access.

2.3 CLIF is starting from a point of agnosticism about the direction content will flow between the repository and other systems (the lifecycle may require movement in both directions). Nevertheless, by facilitating the links between systems it is intended to support preservation by allowing the content to be moved to a system that has preservation capability.

2.4 The development of preservation policies for the repositories as part of the project will guide the technical work proposed. Whilst looking at specific policies for preservation, the potential of incorporating the principles involved into wider institutional policies supporting research, teaching and administration will also be explored, to link the management of the content to the purpose for which it is being managed.

It is expected that the project will produce the following outcomes and value:

- A better understanding of the content lifecycle for different types of content as they are used for the purposes of research, teaching and administration. Understanding this will assist in planning the implementation of systems on an institution-wide basis and facilitate the integration of repositories into institutional environments.

- Documented use cases to support the understanding indicated above, which will provide insights into how content passes through an institution so that it can be managed without 'falling between the cracks'.

- Technical documentation on the integrations carried out, at the three levels indicated in the introduction. By highlighting how systems can work together to support the content lifecycle it is hoped that perceived duplication between systems and the role each could play will be clarified.

- A technical architecture to support the content lifecycle, which demonstrates how systems can be linked together to best support the lifecycle of different types of content. Key to this will be a recognition that not all content will be managed in the same way, but that systems may need to be linked in different ways to meet different needs.

## 3. Overall Approach

3.1 The project will use a methodology that has been tried and tested in past successful projects carried out at the University of Hull. This places the user at the centre of technical development, to ensure that the development that does take place is relevant for the purpose at hand.

3.2 CLIF's first phase will start with range of review work before any implementation takes place.

- A content lifecycle literature review will seek to provide a coherent picture of how a lifecycle approach can be used to support institutional digital content management. The review will be undertaken as cross-disciplinary desk research in liaison with the contributing academics for their subject and role-related input.

- In order to ensure that the implementation is grounded in authentic institutional practices and requirements we will produce detailed use cases in consultation with the academics involved in the project, who represent a variety of processes and types of digital content. The use cases will address issues around the creation, management, processing, annotation, and use/re-use of digital content in a *shared, collaborative* institutional environment, and will inform the development of appropriate policies. Cross-institutional influences on the content lifecycle will also be explored. Evaluation of the functionality provided will focus on the targeted academics, but we will also aim to involve a wider group of stakeholders in trying out the developed solutions (with greater breadth but less depth) and provide feedback.

- Fedora, MOSS and Sakai all provide a rich and complex set of functionalities, thus the design and development work will thus be preceded by a review of the available functionality, with particular regard to the use cases. The ability of each system to support different stages in the content lifecycle, where appropriate integration points are located, and how they will be enabled will also be examined in detail.

- Integrations between systems can be carried out using point-to-point techniques according to specific need. Whilst a loosely coupled approach to point-to-point can enable wider adoption of a solution, such solutions can also be limited by the systems themselves as they change over time. Enterprise Service Buses are an approach to abstract out the ways that systems can communicate with each other, protecting integrations against software changes. This final piece of review work will specifically examine available options for using an ESB-approach to inform subsequent technical development.

3.3 The CLIF team, thus informed, will then develop an overall architecture for the integration of the three systems involved. This will take two forms: a specific architecture to inform the subsequent development phase of the project; and an abstract version of the architecture that addresses the issues involved for readers outside the project. Both will be based around how the content lifecycle can be managed. The architecture will underpin technical work which will be carried out on an iterative basis to allow testing and validation by users.

3.4

## 4. Project Outputs

The project work packages (WPs) will result in the following deliverables:

### *WP1 Project Management*

- Project Plan
- Periodic progress reports
- Final and Completion Reports

### *WP2 Content lifecycle literature review*

- Literature review document

### *WP3 User engagement and policy development*

- Use case documentation

### *WP4 Fedora/Sakai/MOSS integration technical research and investigation*

- Technical review document
- Small prototypes for validating individual issues

### *WP5 ESB feasibility investigation*

- ESB review document

### *WP6 Technical architecture*

- Architecture documentation

### *WP7 Development and implementation*

- Technical specification document
- Software components

### *WP8 Testing*

- Test documentation
- Technical evaluation report

### *WP9 Dissemination*

- Website
- Presentations
- Workshops

### *WP10 Coherence*

- Comprehensive 'About' page on partner repositories

### *WP11 Sustainability*

- Software outputs in regular use with partners and tested elsewhere

## 5. Project Outcomes

The outputs described above will allow the CLIF project to bring about two key outcomes. The first will be a fuller understanding, shared with the community, of the lifecycles that digital, and especially born-digital, materials undertake. The project will then show how Fedora, Sakai and/or MOSS technologies can be brought together to provide a coherent framework in which to provide integrated management of these objects throughout their lifecycles.

## 6. Stakeholder Analysis

| Stakeholder | Interest / stake | Importance |
|---|---|---|
| Lecturers | Lecturers produce many digital materials for their teaching, and often make these available through a VLE. If this content can be placed in a repository it can be better preserved for future use and made available through other environments as well | High |
| Researchers / research management staff | Research datasets are forming an increasing body of digital content. They will be managed for different purposes at different times: having a clear picture of how they are managed throughout their lifecycle will assist the research management process. | High |
| Admin staff / records managers | Lifecycle management of content is key to records managers and they have long looked to enable this. The ability for content to flow between systems will allow administrative decisions to be based on lifecycle need rather than system capability. | High |
| Librarians / Information managers | Librarians will benefit from greater knowledge of how lifecycle management applies in the digital world. Information managers may have specific remits within the overall lifecycle: an awareness of related stages will help place their work in context. | High |
| Computing staff | Many system implementations are either initiated or taken forward by Computing Services staff. Wider awareness of the content lifecycle will assist in scoping system requirements and encourage cross-institutional planning. | Medium / High |
| Software communities | Each of the systems being used has an established community of users. Understanding how a system can integrate with others to facilitate broader usage and increased functionality will support related initiatives that can build on this work. | Medium / Low |
| JISC | The project seeks to embed repositories within institutional environments in a way that focuses on the processes the institutions need rather than the systems themselves and will work with the JISC to develop this approach alongside related projects. | High |
| Institutions with related infrastructure components | Institutions that have similar infrastructural components (repository, VLE, MOSS) will also be engaged through the project to further identify the issues in managing content through its lifecycle. | Medium / High |

## 7. Risk Analysis

| Risk | Probability (1-5) | Severity (1-5) | Score (P x S) | Action to Prevent/Manage Risk |
|---|---|---|---|---|
| **Staffing** | | | | |
| Difficulties recruiting and retaining staff | 2 | 4 | 8 | Most staff involved are in post already. Related staff have wide knowledge of the environment and can contribute |
| **Organisational** | | | | |
| Breakdown in link between partners | 1 | 5 | 5 | Partners know each other. Regular communications to avoid any conflicts. |
| Difficulty in getting required input from academics | 1 | 2 | 2 | The academics have all expressed their interest and commitment (see supporting letters). Alternatives will be identified if problems arise. |
| Failure to meet project milestones | 2 | 3 | 6 | Produce project plan with clear objectives. Follow evaluation plan (see section 4.4). |
| **Technical** | | | | |
| A complete solution cannot be implemented | 2 | 3 | 6 | The absence of a complete solution is not an indication of failure, as there are many potential solutions to parts of the project. The project report would address the issues that could not be resolved. |
| Barriers to integration identified | 3 | 3 | 9 | Work with software communities involved to identify solution or workaround |
| **External** | | | | |
| Lack of engagement from user and software communities | 1 | 4 | 4 | Community contacts have indicated wide interest (see also supporting letters). Make use of personal contacts and community events to raise engagement. |
| **Legal** | | | | |
| Licensing of software outputs affected by use of proprietary software | 2 | 4 | 8 | Identify likelihood of this restriction with Microsoft in early stages. Work with OSSWatch to identify solution. |

## 8. Standards

As far as possible, the work will use open standards at different levels of integration. For moving content potential standards will include OAI-PMH, and OAI-ORE, for structuring content options include METS, IMS CP, ATOM and RAMLET. For interfaces between systems, specifications such as JSR 283 and web service standards will be explored. Fedora now has both SOAP and REST APIs, and we will investigate the relative merits of each. We will follow the principle of loose coupling between Fedora and the wider environments (MOSS, Sakai) with which we will integrate it using clear and well-defined interfaces. As well as being a sound design policy in general, this will make it easier for other institutions to apply/adapt our methodology and/or software modules when integrating with other repository software or other institutional environments. As one of the aims of the project is to integrate Fedora with MOSS, it is inevitable that some development will involve proprietary Microsoft approaches, in particular the use of .NET, C# and WebParts. However, we will endeavour to isolate such instances, making use of open standards, wherever feasible, such as WSRP and/or BPEL.

| Name of standard or specification | Version | Notes |
|---|---|---|
| OAI-PMH | 2.0 | OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is a protocol developed by the Open Archives Initiative. It is used to harvest (or collect) the metadata descriptions of the records in an archive so that services can be built using metadata from many archives. |
| OAI-ORE | 1.0 | Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of Web resources. The OAI-ORE specifications are based around the ORE Model which introduces the Resource Map (ReM) that makes it possible to associate an identity with aggregations of resources and make assertions about their structure and semantics. |
| METS | 1.7 | The Metadata Encoding and Transmission Standard schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. |

| | | |
|---|---|---|
| IMS CP | 1.1.4 or 1.2 (tbd) | The IMS Content Packaging Specification provides the functionality to describe and package learning materials, such as an individual course or a collection of courses, into interoperable, distributable packages. Content Packaging addresses the description, structure, and location of online learning materials and the definition of some particular content types. |
| ATOM | 1.0 | ATOM is a syndication format allowing focussed information delivery |
| RAMLET | | RAMLET is developing a standard that will help understand, represent and map diverse resource aggregation specifications and standards. In simple terms, it provides a framework for mapping different types of packaging specifications to each other. |
| JSR 283 | 2 | Content Repository API for Java is a specification for a Java platform API for accessing content repositories in a uniform manner. The content repositories are used in content management systems to keep the content data and also the meta-data used in CMS such as versioning meta-data. The specification was developed under the Java Community Process as JSR-170 (Version 1) and as JSR-283 (Version 2). |
| WSRP | 2.0 | WSRP defines a set of interfaces and related semantics which standardise interactions with components providing user-facing markup, including the processing of user interactions with that markup. This allows applications to consume such components as providing a portion of the overall user application without having to write unique code for interacting with each component. |
| BPEL | 2.0 | (WS)BPEL is a standard originating from IBM designed to orchestrate Web services so as to support business transactions. |

## 9. Technical Development

To ensure that the software and functionality produced meet the needs of the targeted users and other stakeholders within the institution the project will follow a user-driven, evolutionary approach, involving incremental cycles of development, implementation and evaluation in collaboration with these users and stakeholders.

## 10. Intellectual Property Rights

IPR in all reports and other documents produced by project staff as part of the project will be retained by the authors and host institution but made freely available on a non-exclusive licence as required by JISC. We will respect the licence model of all third party software used during the project. Noting this, the project will look to make any software created during the project available to the community on an OSI-approved open-source basis, taking advice from OSSWatch as required.

# Project Resources

## 11. Project Partners

### *11.1 University of Hull*

Project management, requirements gathering, workflow development and user testing, technical development and investigations, dissemination and evaluation activities.

Contact: Chris Awre, Project Director (c.awre@hull.ac.uk)

#### *11.1.1 Richard Green, IT Consultant (subcontractor)*

Project management and coordination on behalf of the University of Hull.

Contact: Richard Green, Project Manager (r.green@hull.ac.uk)

**Note: This is the main project contact**

### *11.2 Centre for e-Research, King's College London*

Requirements gathering, workflow development and user testing, technical development and investigations, dissemination and evaluation activities.

Contact: Mark Hedges, Site Manager for KCL work (mark.hedges@kcl.ac.uk)

### *11.3 Consortium Agreement*

The Consortium Agreement was signed on *tba.*

## 12. Project Management

### *12.1. University of Hull*

| | | |
|---|---|---|
| Chris Awre, Project Director | c.awre@hull.ac.uk | |
| Richard Green, Project Manager(0.5 fte) | r.green@hull.ac.uk | (subcontractor) |
| Analyst/developer | to be appointed | |
| Vicky Mays, University Records Manager | v.mays@hull.ac.uk | |
| Warren Viant, Head of Computer Science | w.j.viant@hull.ac.uk | |
| Eur Ing Tony Wilkinson, Dept of Engineering | a.j.wilkinson@hull.ac.uk | |

**Chris Awre** will act as Project Director, overseeing the project as a whole in the context of wider institutional, technical and JISC-related issues. **Richard Green** will act as Project Manager, overseeing the day-to-day activity of the project across all workpackages, and working closely with the site manager at CeRch. **Vicky Mays** will act as liaison between the project and records management use cases. **Warren Viant** and his departmental colleagues have expertise in the implementation of MOSS and use Fedora for an undergraduate dissertation collection; thus he can contribute to both use cases and technical advice. **Eur Ing Tony Wilkinson** has an interest and need in managing student report and marked work submissions, lab sheets, program source files,

programs to download, and application data files, and is an active user of Sakai for teaching. He too will contribute to both use cases and technical advice.

### *12.2. King's College London*

| | |
|---|---|
| Dr Mark Hedges, Site Manager | mark.hedges@kcl.ac.uk |
| Analyst/developer | to be appointed |
| Dr David Betz, Department of War Studies | david.betz@kcl.ac.uk |
| Dr Stella Fabiane, Randall Division of Cell and Molecular Biophysics | stella.fabiane@kcl.ac.uk |
| Dr Mark Mulligan, Department of Geography | mark.mulligan@kcl.ac.uk |

**Dr Mark Hedges** will act as site manager for the project at KCL and will oversee work there. **Dr David Betz** uses resources which include documents, reports, speeches and texts, as well as interviews, which are mostly informal and unstructured. The majority (99%) of this material is digital, and may include audio and visual material. He will contribute to CLIF's use cases. **Dr Stella Fabiane** is a researcher whose research outputs include not only traditional journals, but also the processed data sets. She too will contribute to our use cases. **Dr Mark Mulligan** maintains a set of very large environmental datasets which will provide our final set of use cases.

## 13. Programme Support

The project team would be grateful if the JISC would:

- provide adequate advanced notice of programme meetings and non-standard reporting requirements
- identify potential areas of collaboration or communication with projects in other programmes

## 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 |
|---|---|---|---|---|
| At project management meetings | Cross-partner working effectiveness | Is work progressing as expected across both partners | Agenda item for meetings | Noted that all is OK or action to address if not |
| End months 4, 8, 12, 16 | Stakeholder engagement and feedback | Are users engaged? What are they feeding back? | Record feedback and assess | Increasing engagement over time. Input for the project. |
| End month 6 | Outcomes of WP4 and WP5 | What technical approaches are most valid to pursue? | Project meeting | Agreement of how WP7 is to be taken forward |
| Monthly during development | Quality of code outputs | Is code stable? | Code review. JUnit testing | Successful testing and review |
| Monthly during testing | Progress of testing | Is testing providing what we need? | Review of testing with those involved | Acknowledged progress in testing |
| End month 19 | Whether use cases have been met | Have use cases been successfully supported? | Review against use cases from WP3 | Agreement with departmental partners of successful outcome |

Gathering formative evaluation from stakeholders is formally scheduled in the above table, though the project will seek to make the process continuous and use the schedule as a series of hooks for assessing progress. This schedule is built around the beginning and end of appropriate workpackages on which to gather feedback.

## 17. Quality Plan

| **Output** | | | | | |
|---|---|---|---|---|---|
| **Timing** | **Quality criteria** | **QA method(s)** | **Evidence of compliance** | **Quality responsibilities** | **Quality tools (if applicable)** |
| **D1: Literature review document** | | | | | |
| 04/09 | Meets project and community needs | Community peer review | Community feedback | RG with CA | |

| | | | | | |
|---|---|---|---|---|---|
| **D2: Use case documentation** | | | | | |
| 10/09 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D3: Technical review document** | | | | | |
| 10/09 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D4: ESB review document** | | | | | |
| 10/09 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D5: Architecture documentation** | | | | | |
| 11/09 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D6: Technical specification document** | | | | | |
| 10/10 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D7: Test documentation** | | | | | |
| 12/10 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH, developer/ analysts | Yet to be agreed but to include JUnit |
| **D8: Technical evaluation report** | | | | | |
| 12/10 | FFP | Peer review* and comment from Programme Manager | Acceptance by reviewers | RG, CA, MH | |
| **D11: Sustainability** | | | | | |
| 12/10 | FFP | Software tested by non-partner institutions | Acceptance by testers | RG, CA, MH, developers | |

FFP= fit for purpose

* These documents will be shared with others in the field with appropriate domain knowledge and feedback obtained. In particular we shall use colleagues at the University of Virginia and Stanford University who have a particular interest in following this work.

## 18. Dissemination Plan

| Dissemination Activities | Audiences | Purpose | Key Messages |
|---|---|---|---|
| - Conference presentations and workshops<br>- Demonstrators (including testing by others)<br>- Website<br>- Availability of outputs within a repository<br>- Mailings to lists | - Stakeholders<br>- Attendees at meetings organised by communication routes listed above<br>- Programme meetings, IEEE Science meetings, Internet Librarian, Online, Open Repositories | To engage the relevant user communities and demonstrate the validity and benefits of the work | - The role and benefits of a content lifecycle approach to managing digital content<br>- How project outputs can be exploited by their institutions<br>- Practical relevance, demonstrated by the use cases<br>- Pragmatic approach - concrete outputs but loosely coupled for adaptation<br>- Commitment of partners to using the systems involved |

## 19. Exit and Sustainability Plans

| Project Outputs | Action for Take-up & Embedding | Action for Exit |
|---|---|---|
| Documentation | Dissemination through conferences and workshops (see dissemination plan). Summaries presented through blog posts. | Ongoing availability through website and repository (-ies) (JISC and institutional) |
| Software | Testing within partners and through other sites (see stakeholder analysis). Demonstration of software working in situ. | Ongoing availability of software components through website and Sourceforge or similar site |
| Pre-production implementations | Implementation of working pre-production environments at the institutions | Use the environments established to encourage engagement by other users in the institutions |
| Website | Announcements through mailing lists | Maintenance of website for 3 years, plus contribution to web archive |
| Presentations | Given at conferences (see dissemination plan) | Ongoing availability of presentations on website and via Slideshare |

| Project Outputs | Why Sustainable | Scenarios for Taking Forward | Issues to Address |
|---|---|---|---|
| Sustainable software outputs in regular use | Fit for purpose, fulfilling a community need | Available for download | Appropriate download facilities |

# Appendices

## Appendix A1: Full budget

## Appendix A2: Project Budget (JISC contributions only)

## Appendix B: Workpackages

**CLIF Project 1st April 2009 - 31st December 2010**

| | | | 2009 | | | | | | | | | 2010 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | A | M | J | J | A | S | O | N | D | J | F | M | A | M | J | J | A | S | O | N | D |
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| **WP1** | **Project management** | | | | | | | | | | | | | | | | | | | | | | |
| | | Project plan | | | | | | | | | | | | | | | | | | | | | |
| | | Progress reports | | | | | | | | | | | | | | | | | | | | | |
| | | Final and completion reports | | | | | | | | | | | | | | | | | | | | | |
| **WP2** | **Content lifecycle literature review** | | | | | | | | | | | | | | | | | | | | | | |
| | D1 | Literature review document | | | | | D1 | | | | | | | | | | | | | | | | |
| **WP3** | **User engagement and policy development** | | | | | | | | | | | | | | | | | | | | | | |
| | D2 | Use case documentation | | | | | | | D2 | | | | | | | | | | | | | | D2 |
| **WP4** | **Fedora/Sakai/MOSS integration research & investigation** | | | | | | | | | | | | | | | | | | | | | | |
| | D3 | Technical review document | | | | | | | D3 | | | | | | | | | | | | | | |
| **WP5** | **ESB feasibility investigation** | | | | | | | | | | | | | | | | | | | | | | |
| | D4 | ESB review document | | | | | | | D4 | | | | | | | | | | | | | | |
| **WP6** | **Technical architecture** | | | | | | | | | | | | | | | | | | | | | | |
| | D5 | Architecture documentation | | | | | | | | D5 | | | | | | | | | | | | | |
| **WP7** | **Development and implementation** | | | | | | | | | | | | | | | | | | | | | | |
| | D6 | Technical specification document | | | | | | | | | | | | | | | | | | | D6 | | |
| **WP8** | **Testing** | | | | | | | | | | | | | | | | | | | | | | |
| | D7 | Test documentation | | | | | | | | | | | | | | | | | | | | | D7 |
| | D8 | Technical evaluation report | | | | | | | | | | | | | | | | | | | | | D8 |
| **WP9** | **Dissemination** | | | | | | | | | | | | | | | | | | | | | | |
| | | Website | | | | | | | | | | | | | | | | | | | | | |
| | | Presentations and workshops | | | | | | | | | | | | | | | | | | | | | |
| **WP10** | **Coherence** | | | | | | | | | | | | | | | | | | | | | | |
| | | Comprehensive 'about' information on repository | | | | | | | | | | | | | | | | | | | | | |
| **WP11** | **Exit and sustainability** | | | | | | | | | | | | | | | | | | | | | | |

| Workpackage and activity | Earliest start date | Latest completion date | Outputs | Milestone | Responsibility |
|---|---|---|---|---|---|
| **YEAR 1** | | | | | |
| ***WP1: Project management***<br>Management of the work including: planning; coordination of contributing academics; monitoring progress on technical work; advocacy. | 01/04/09 | 31/12/10 | | | RG, CA, MH |
| 1. Interim project report 1 | | | **Progress report** | 31/10/09 | RG |
| 2. Interim project report 2 | | | **Progress report** | 30/04/10 | RG |
| 3. Interim project report 3 | | | **Progress report** | 31/10/10 | RG |
| 4. Final and completion reports | | | **Final and completion reports** | 31/12/10 | RG, CA, MH |
| ***WP2: Content lifecycle literature review***<br>This workpackage will seek to provide a coherent picture of how a lifecycle approach can be used to support institutional digital content management. The literature review will be undertaken as cross-disciplinary desk research in liaison with the contributing academics for their subject and role-related input. | 01/05/09 | 31/07/09 | | | CA, academics |
| 5. | | | **Literature review document** | 31/07/09 | CA |

| | | | | | |
|---|---|---|---|---|---|
| ***WP3: User engagement and policy development***<br>To ensure that the implementation is grounded in authentic institutional practices and requirements we will produce detailed use cases in consultation with the academics involved in the project, who represent a variety of processes and types of digital content. The use cases will address issues around the creation, management, processing, annotation, and use/re-use of digital content in a *shared, collaborative* institutional environment, and will inform the development of appropriate policies. Cross-institutional influences on the content lifecycle will also be explored. Evaluation of the functionality provided will focus on the targeted academics, but we will also aim to involve a wider group of stakeholders in trying out the developed solutions (with greater breadth but less depth) and provide feedback. | 01/06/09 | 30/09/09 | | | *Omnes* |
| 6. | | | **Use case documentation** | 30/09/09 | RG, CA, MH |
| ***WP4: Fedora/Sakai/MOSS integration research and evaluation***<br>Fedora, MOSS and Sakai all provide a rich and complex set of functionalities. The project team has differing levels of experience of using the three systems. All design and development work will thus be preceded by a review of the available functionality, with particular regard to the use cases. The ability of each system to support different stages in the content lifecycle, where appropriate integration points are located, and how they will be enabled will also be examined in detail. | 01/04/09 | 30/09/09 | | | RG, CA, MH, WV, TW |
| 7. | | | **Technical review document** | 30/09/09 | RG, CA, MH |

| | | | | | |
|---|---|---|---|---|---|
| ***WP5: ESB feasibility investigation***<br>Integrations between systems can be carried out using point-to-point techniques according to specific need. Whilst a loosely coupled approach to point-to-point can enable wider adoption of a solution, such solutions can also be limited by the systems themselves as they change over time. Enterprise Service Buses are an approach to abstract out the ways that systems can communicate with each other, protecting integrations against software changes. This workpackage will specifically examine available options for using an ESB-approach to inform subsequent technical development. | 01/07/09 | 30/09/09 | | | RG, CA, MH |
| 8. | | | **ESB review document** | 30/09/09 | RG, CA, MH |
| ***WP6: Technical architecture***<br>Following on from WP4 and WP5, an overall architecture for the integration of the three systems involved will be developed. This will take two forms: a specific architecture to inform the subsequent development phase of the project; and an abstract version of the architecture that addresses the issues involved for readers outside the project. Both will be based around how the content lifecycle can be managed. | 01/09/09 | 31/10/09 | | | RG, CA, MH, developer/ analysts |
| 9. | | | **Architecture documentation** | | RG, CA, MH |

| | | | | | |
|---|---|---|---|---|---|
| ***WP7: Development and implementation***<br>This workpackage covers the detailed design, coding, unit testing and integration of the software components. An iterative approach will allow this workpackage to overlap chronologically with WP 8 and 9 in order to maximise the available time and maintain communication with those testing the developments. | 01/10/09 | 30/09/10 | | | RG, CA, MH, developer/ analysts |
| 10. | | | **Technical specification** | 30/09/10 | RG, CA, MH |
| ***WP8: Testing***<br>We intend to produce software that is usable in a production environment, at least as a "perpetual beta": rigorous and documented testing is thus of key importance. Given the incremental approach to development, testing may be repetitive, so where possible we will create sets of automated test scenarios that can be executed against the software without human interaction, to validate the software after any significant changes have been made. For consistency, testing will take place in a common environment whilst taking into account specific local differences. | 01/02/10 | 30/11/10 | | | RG, CA, MH, developer/ analysts |
| 11. | | | **Test documentation** | | Developer/ analysts |
| 12. | | | **Technical evaluation reports** | | RG, CA, MH |

| | | | | | |
|---|---|---|---|---|---|
| ***WP9: Dissemination***<br>The dissemination workpackage will focus activity on communicating the ideas and outputs of the project to a range of audiences. The work will be coordinated and carried out by both partners as opportunity arises. | 01/11/09 | 31/12/10 | | | *Omnes* |
| 13. | | | Website | 30/04/09 | RG |
| 14. | | | Presentations and workshops | | *Omnes* |
| ***WP10: Coherence***<br>The briefing document accompanying the 12/08 call identified six activities that should be undertaken to address coherence across repositories. These will be addressed by both partners during the project such that an 'about' page dealing with them can be provided in their repositories. | | | | | |
| 15. | 01/07/09 | 31/12/10 | 'About' page | 31/12/10 | RG, CA, MH |
| ***WP11: Exit and sustainability***<br>This workpackage will focus activity on how the project's work can be used beyond the lifetime of the project. Key to this will be the partners' commitment to implementing the solution locally, testing it elsewhere, and demonstrating it working in practice. | 01/01/10 | 31/12/10 | | | *Omnes* |
| 16. | | | Sustainable software outputs in regular use | 31/12/10 | *Omnes* |

## References

[1] See Appendix B for examples of content lifecycles

[2] JISC Supporting Institutional Records Management programme, http://www.jisc.ac.uk/whatwedo/programmes/supportingirm

[3] Study of the Records Lifecycle project, http://www.webarchive.org.uk/pan/13734/20060324/www.jisc.ac.uk/index79bf.html?name=recordsman_papers_cycle

[4] MoReq2 specification, http://www.moreq2.eu/index.htm

[5] For example, http://www.ecmconnection.com/article.mvc/Learn-From-Content-Lifecycle-Transformation-0002

[6] See http://it.toolbox.com/blogs/pcm/standards-and-content-lifecycle-18174

[7] DCC Curation Lifecycle Model, http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf

[8] Pennock, M. Digital curation: a life-cycle approach to managing and preserving usable digital information, Library & Archives Journal, 2007, Issue 1, http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf

[9] UIUC Digital Content Lifecycle Management project, http://www.library.uiuc.edu/nsm/digcon/

[10] LIFE project, http://www.life.ac.uk/

[11] See http://en.wikipedia.org/wiki/Information_lifecycle_management