

Project Acronym: CTREP  
 Version:  
 Contact: I E Boston  
 Date: 18/06/2007



## Project Document Cover Sheet

Project Information			
<b>Project Acronym</b>	CTREP		
<b>Project Title</b>	Cambridge Tetra Repositories Enhancement Project		
<b>Start Date</b>	01/03/2007	<b>End Date</b>	31/03/2009
<b>Lead Institution</b>	University of Cambridge		
<b>Project Director</b>	John Norman		
<b>Project Manager &amp; contact details</b>	Ian Boston, ian@caret.cam.ac.uk		
<b>Partner Institutions</b>	University of Highlands and Islands, MIT Digital Libraries		
<b>Project Web URL</b>	<a href="https://camtools.caret.cam.ac.uk/access/wiki/site/jisc-ctrep/home.html">https://camtools.caret.cam.ac.uk/access/wiki/site/jisc-ctrep/home.html</a>		
<b>Programme Name (and number)</b>	<i>Capital Programme Sept 06</i>		
<b>Programme Manager</b>	Amber Thomas		

Document Name			
<b>Document Title</b>	<i>Project Plan</i>		
<b>Reporting Period</b>			
<b>Author(s) &amp; project role</b>	Ian Boston, Project Manager		
<b>Date</b>	18/06/2007	<b>Filename</b>	ProjectPlan20070618.doc
<b>URL</b>	TBC		
<b>Access</b>	o Project and JISC internal		o General dissemination

Document History		
Version	Date	Comments
1.0	18/06/2007	Initial Version



## JISC Project Plan

### *Overview of Project*

#### 1. Background

1. The University of Cambridge Information Strategy includes the adoption and support of a University-wide digital repository to collect, preserve and disseminate the University's intellectual assets that are manifest as digital materials. The chosen platform is DSpace, which implements the open standards of the JISC Information Environment. This proposal includes plans to develop the infrastructure and services of the digital repository project to encourage greater use of the repository, to the benefit of the University and other participants in the information environment who will be able to access resources from the repository. Widespread adoption and integration into daily practice is seen as the surest route to embedding and sustainability of the repository. To encourage adoption, tools will be developed that address specific uptake issues at Cambridge, but which are expected to benefit many other institutions adopting digital repository solutions. New tools will be released under open-source licences and will be designed to be compatible with and to contribute to the JISC e-Framework. Full engagement with the e-Framework project is anticipated.
2. A primary output of the project is a seamless integration of repository deposit into the daily research, learning and teaching workflows of faculty. This is expected to lead to a step change in the volume of material deposited as barriers to deposit are lowered. We also expect that the quality and relevance of materials on deposit to be improved by deposit workflows that enforce agreed policies. The negotiation and enforcement of collection and preservation policies is expected to permit responsible management of the preservation commitment as the volume of material on deposit grows.
3. The University of Cambridge institutional repository will provide an institutional open access repository of research outputs as well as a repository of teaching and learning materials. A key output of this project will be integration with the University collaboration environment for research groups and the course management system (CamTools<sup>1</sup>). The approach taken will be generalised and broadly applicable to other institutions and other platforms in compliance with the goals of the JISC/DEST e-Framework.
4. The UHI is committed to adoption of the TETRA Framework as a part of the TETRA Collaboration. Within that strategy is a need to migrate to a fully integrated institutional repository, based on open standards and open architectures that enhance collaboration amongst the member organisations of UHI.
5. This is a repository enhancement project in which the consortium is constructed to ensure maximum value and transferability of the software outputs and lessons learned to the wider JISC community. To ensure that these outputs are applicable to the community, UHI will be porting to a Fedora based repository and deploying a pilot Fedora IR at UHI. Consequently the distribution of workload to achieve significantly increased take up is not symmetrical.
6. Research projects have already been initiated that have negotiated data archiving strategies with the AHDS by deposit in the University DSpace repository.
7. The University is also engaged in a subject-based repository (TLRP for Education Research) that addresses an identified need across institutions using the same DSpace software. The repository

1. \_\_\_\_\_  
<sup>1</sup> <https://camtools.caret.cam.ac.uk/portal>

caters for a broad range of content. It is anticipated that the tools created for the institutional repository will be immediately useful for the purposes of this subject-based research community.

8. Cambridge is fully committed to the adoption and use of open-standards and contributed to the JSR168 development in the RDN portal project (now Intute).

9. As a founding partner of the Sakai project Caret has driven the thinking on the integration between Content Repositories and Institutional Repositories in the Sakai project. This work has included JSR-170 implementation within Sakai and joint presentations and white papers with the Digital Libraries group at MIT.<sup>2</sup>

10. UHI and University of Cambridge are founding members of the TETRA Collaboration and will ensure that the dissemination of repository policy between a Fedora instance and a future TETRA Framework is supported. This work will benefit the other TETRA Collaboration members namely University of Oxford and University of Hull<sup>3</sup>.

11. The project will build on the work undertaken by MIT in the PLEDGE project to express policies. The Sakai integration will be performed by building on the work undertaken by CARET to integrate repositories in general within Sakai Content Hosting Service, and deploy this feature into CamTools to use the Deposit API based on DSpace LNI from the group at UKOLN, the strand of this work being undertaken by Jim Downing at Cambridge.

12. UHI currently maintains an Intralibrary DLOR and has been running a project to embed this repository into institutional process for a couple of years with limited success. It is hoped that the integration opportunities found within CTREP will help to meet this objective within UHI and make DR use a more normalized business process. Furthermore, UHI has been taking part in the JISC funded Trust DR project, which is examining the practical issues in setting up the DRM in LORs.  
<http://trustdr.ulster.ac.uk>, <http://cms.uhi.ac.uk>

## 2. Aims and Objectives

13. The overall project falls into 4 parts:
- Secure long-term institutional commitment (done)
  - Establish a secure and scalable technical infrastructure (in progress)
    - Employ 2 staff to administer hardware and software and to resolve technical issues on ingest, search and retrieval, and OAI interfaces
  - Communicate, educate and support adopters
    - Employ project manager to coordinate initiatives and manage internal and external relationships
    - Employ support staff to communicate, educate and support adopters
  - Develop tools and processes to address barriers to use and to enhance interoperability with national projects
    - Develop deposit API
    - Develop policy expression<sup>4</sup> for collections and policy enforcement in online workspaces for individuals, research groups and courses and integrate into natural workflows
14. The main development focus of the project is to:
- Integrate CamTools with DSpace at Cambridge Institutional Repository such that:
    - The policies of the IR are understood by CamTools
    - Material prepared for deposit in the IR is cleared by a pre ingest workflow managed by the users of the CamTools environment

1. \_\_\_\_\_

<sup>2</sup> <http://cwspace.mit.edu/docs/ProjectMgt/Reports/SakaiVancouver2006/REILLY-DSpaceSakaiVancouver2006.ppt>

<sup>3</sup> [http://www.elearning.ac.uk/news\\_folder/tetracollaboration](http://www.elearning.ac.uk/news_folder/tetracollaboration)

<sup>4</sup> <http://libraries.mit.edu/dlrg/projects.html#pledge>

- Every Personal and Group workspace may have an integrated area of the Resources tool (file manager) that represents an area within the IR.
  - In parallel extend the implementation to work with Fedora
  - Deploy this integration to a selection of communities in Cambridge
  - Evaluate the impact of this deployment on the DSpace at Cambridge project with a view to implementing ingest workflows designed at managing that workload or iteratively improving the deposit policy statements.
  - Integrate a 3<sup>rd</sup> party Activity based Workflow Management<sup>5</sup> system with DSpace
  - Implement improved ingest Workflow and reviewed policies
  - Evaluate impact of improved ingest workflow with a view to continuing improvement by measurement after the end of the project.
15. Resulting in:
- An integration between the IR and VLR/VRE that supports the actively managed deposit and ingest of teaching learning and research materials with improved meta-data quality, streamlined ingest and wider reuse.
  - To perform the above tasks in such a way as to ensure that the components can be used by TETRA Collaboration.

### 3. Overall Approach

16. The Project will be structured into 2 phases representing two cycles of enhancement. The Outcomes of the evaluation performed in phase 1 will inform the development of systems in phase 2 aiming to deliver yearly cycles of continuous improvement based on clearly acquired metrics. Since this project will be deploying into production environments the development and deployment cycles will be synchronised with the production environment at Cambridge, hence the yearly cycles. To achieve this the project cycle will be using agile open source development strategies embodied in methodologies like SCRUM<sup>6</sup>, delivering bi-weekly cycles of development.

17. The project will use CamTools (Cambridge VLE/VRE implementation) and the DSpace institutional Repository. There will be a pilot implementation of Fedora with the TETRA Framework at UHI. This architecture stack demands that existing Standards are used over network interfaces. By involving two institutions and a variety of Repositories we will ensure that the components produced adopt Standards based integration and support the JISC Information Environment. Integration between CamTools and DSpace will use the DSpace LNI interface that is a WebDAV based protocol. Integration between the TETRA Framework and Fedora intends to use an OKI-DR OSID, as currently being used by other Sakai partners. We will also look at using the evolving Deposit API being developed by UKOLN and Jim Downing's work with a WebDAV based Deposit API implementation.

18. The main focus of the development will be on the integration of Repository Policy expression into the dissemination and deposit connection between the VLE/VRE and the Institutional Repository.

19. The Project will be informed by the guidance given by SHERPA in the areas of Metadata and preservation policy, though the DSpace@Cambridge partnership membership of the SHERPA and the SHERPA officer at Cambridge.<sup>7</sup>

20. The project will be informed by the work and outputs of the RepMMan project at Hull, which is also part of the TETRA Collaboration.

21. It is expected that the evaluation undertaken in phase 1 of the project will indicate that the impact of wider access to the institutional repository will increase the deposit rate as well as the quality of the deposits. If the project achieves this phase 1 will be a success. In phase 2 we will apply workflow techniques in the deposit process and in the ingest process to reduce the manual processing associated with a greater uptake. If the evaluation at the end of phase 2 shows that this work has resulted in a positive impact, then phase 2 will also be a success.

1. \_\_\_\_\_

<sup>5</sup> Activity Based Workflow based on the same principals as RAMS and WfMC.

[http://www.lamscommunity.org/dotlrn/clubs/educationalcommunity/lamsresearchdevelopment/forums/message-view?message\\_id=318164](http://www.lamscommunity.org/dotlrn/clubs/educationalcommunity/lamsresearchdevelopment/forums/message-view?message_id=318164)

<http://www.wfmc.org/>

<sup>6</sup> [http://en.wikipedia.org/wiki/Scrum\\_\(in\\_management\)](http://en.wikipedia.org/wiki/Scrum_(in_management))

<sup>7</sup> <http://www.sherpa.ac.uk/contacts.html>

22. The project will deliver a sustainable implementation for Cambridge and a prospective implementation for the TETRA Collaboration so that the implementation can be used in a number of other UK HE institutions.

#### 4. Project Outputs

23. Tangible Outputs:

- A step change in the uptake of the DSpace@Cambridge within Cambridge and the potential uptake of Fedora at UHI as the TETRA Framework is adopted.
- Greater availability of research outputs from projects hosted at Cambridge via DSpace@Cambridge open access policy.
- A Sakai Content Hosting Handler<sup>8</sup> implementation for DSpace capable of read and write access to the DSpace repository, with deposit managed by policy expressions informed by the PLEDGE project at MIT.
- An OKI-DR OSID implementation of the same Content Hosting Handler for connection to Fedora
- Deployment of both of these implementations in the institutional VLE/VRE (CamTools) at Cambridge connected to DSpace and a pilot Fedora implementation.
- Publication of sample Policies in use for preservation and deposit
- An online report
  - outlining the results of the above implementation in use in a selected community of research and teaching.
  - indicating the impact of wider participation in repository deposit with recommendations of how these might be addressed
- An implementation of Activity based workflow using an existing workflow engine to enhance DSpace ingest workflow
- An online report on the results of an evaluation focusing on the impact of improved automated ingest workflow on an Institutional Repository.

24. Intangible Outputs:

- A greater understanding in the TETRA Collaboration of the expression of Policy by the Institutional Repository and the effect on its users.
- A building block for use within the TETRA Framework that will enable wider repository integration, use and enhancement amongst institutions deploying the TETRA Framework.
- Fostering wider Open Source collaboration and re-use within UK HE.

#### 5. Project Outcomes

25. Ideally the project will open the Cambridge Institutional Repository up to a wider community, and increase the volume and variety of items being deposited without increasing the resource requirements to curate the repository. This will improve dissemination of items outside Cambridge hopefully including some significant image collections.

26. With a richer collection in the institutional repository and easier access to those items with understood policy, it will become possible to re-use digital assets within research teaching and learning.

27. By involving the TETRA Collaboration in the core implementation of the project, the outputs of this project will be usable in at least 3 other UK HE institutions and available to all UK HE and FE institutions as freely available open source components. Additionally, there will be the knowledge base which will be developed during the project which will be a valuable resource for other interested institutions. This will ensure the skills and techniques developed in this project become available to the HE sector in general to help with the wider maintenance and use of digital assets.

#### 6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
-------------	------------------	------------

1. \_\_\_\_\_  
8 <http://article.gmane.org/gmane.comp.cms.sakai.devel/7851>

University of Cambridge	Interest: Enhancement of its Institutional Repository with improved access, richer content and wider adoption leading to a potential for wider dissemination of the content contained in the IR. Stake: The outputs of the project will be in production at Cambridge supporting the campus.	High
UHI	Interest: Integration of Fedora into TETRA Framework to support future development under TETRA Collaboration. Increase integration and use of IR within existing community. Stake: Committed to migrating to the TETRA Framework, integrating a Sakai instance and a pilot Fedora Repository within the timeframe of the project.	High
University of Cambridge Library	Interest: Better management and dissemination of its preservation and deposit policies with the potential to disseminate other policies. Improved ingest workflow and higher quality of items within the repository without a corresponding increase in resources. Stake: DSpace@Cambridge Project and providing Repository manager and developer resource to the project.	High
TETRA Collaboration	Interest: A mechanism to integrate any standards based Repository into the TETRA Framework with policy expression. Also, the ability to run multiple repositories transparently integrated with a single application container. Finally, with planned SAML 2 AAAS in Sakai using GX, the ability for federated repository sharing between institutions. Stake: Supporting 4 UK HE institutions with Repository needs	High
CARET	Interest: Better integration and between support of the Teaching, Learning and Research communities using CamTools, giving those users the ability to preserve their research data and teaching materials with high quality meta data whilst being informed about the policies under which the preservation is taking place. Stake: Managing CamTools and providing developer resource, evaluation and project management.	High
DSpace Libraries at MIT	Interest: Wider use of the outputs of the PLEDGE project. Possible reuse of the Sakai-DSpace components within MIT institutional deployments such as Stellar Images. Stake: Leading the DSpace project, providing consultancy to the project.	High
JISC	Interest: A two stage evaluation project delivering repository enhancement with clear metrics on the impact of the approach that could be reused elsewhere in UK HE/FE and will inform the development of the update of national repository initiatives, including inter-repository policy expressions. Stake: Funding Institutional Repository enhancement.	Med

## 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 hire suitable staff in a	1	5	5	Although Cambridge has pay and grading in place and this increases the timescales necessary to hire staff, the posts required have already been taken through the pay and

suitable timescale				grading process, and the hiring process is already underway.
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. We make heavy use of collaboration technologies to eliminate travel time and improve co-ordination between partners. Within Cambridge distances are short, but we often use Video Conferencing and Interactive online meeting support (Breeze Meeting) to improve distance collaboration without travel time. The project plan is structured in 2 phases with kick-off meetings and review meetings. The project will re-use existing open source collaboration infrastructure in place at CARET which is on a daily basis in a global open source community.
Technical – technical difficulties prevent implementation and deployment of the technologies that have been selected.	2	2	4	The project will integrate existing technologies where the collaborators already have experience within the Field. CARET has worked closely with the Sakai project and knows the code base intimately. CARET has expertise with Fedora having used it as the underlying repository in the QUADS project. The University Library via DSpace@Cambridge runs a large DSpace implementation and has been closely involved with DSpace for many years. MIT has an immense knowledge base in this area.
External suppliers – the risk of external suppliers being unable to deliver	5	1	5	The project uses generic commercial components that are already in existence, there is no dependence on external suppliers. Where production systems are being used by the project, these have been deployed as highly available clusters to deliver the level of production service required. Since the project will use existing production systems, the risk of issues in this area are already managed and funded by teams outside this project.
Legal – inability to define or resolve the legal issues effecting collaboration	1	5	5	The success of the project and its wider dissemination to UK HE requires that the collaborators do not have legal issues between the institutions that affect the project. As is normal with collaborative projects, agreements have been reached prior to submission of this proposal that cover these areas. This minimises the probability of this risk materialising, but does not entirely eliminate it or reduce its impact.
Legal – inability to define or reach agreement on the legal issues regarding the repositories.	1	5	5	The project is structured to ensure that the institutional deployments do not cross-institutional boundaries minimising potential issues arising from 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.

## 8. Standards

28. Where standards exist they will be used, all communications between distributed systems will use HTTP1.1 and encoding with use REST or SOAP based approaches. Application level protocols exist such as WebDAV these will be used. At a higher level the Deposit API will be used above DSpace LNI. Workflow descriptions will be encoded in BPEL or XPDL, dependant on context. The project will monitor other projects funded by JISC and take account of preferred standards as they develop and are adopted.

Name of standard or	Version	Notes
---------------------	---------	-------

<b>specification</b>		
HTTP	1.1	All network communications will use HTTP 1.1 wherever possible
REST	NA	We will use REST based services where possible.
SOAP, and WSDL	NA	Where there are SOAP and WSDL web services defined these will be used
Deposit API, DSpace LNI	DSpace LNI 1.0	Where a standard deposit API already exists, it will be used eg DSpace LNI. Where no standard exists we will be informed by the work undertaken by SHERPA on Deposit API's
JMS, ESB	NA	For the synchronous connections between applications we will use Java Message Service based connections and may encapsulate these within an ESB
JSR-208, JBI	NA	We expect to use standards based integration points where ever possible. Depending on the choice of integrating infrastructure we may use the Java Business Integration Standard JSR-208.
BPEL, XPD L	BPEL 1.1, XPD L 2.0	For the expression of machine workflows we will use BPEL, for the expression of Activity Base Workflows we will use XPD L. This choice will inform the selection of the Workflow engine for phase 2.
Repositories, DSpace, Fedora	DSpace 1.4.1, Fedora 2.2	We will use DSpace 1.4.1 and Fedora as part of the deployed IR infrastructure.
Java 1.5, Maven, SVN JIRA		We will use Java 1.5 for coding, Maven 1 or 2 for build infrastructure. Code will be in a public SVN repository and we will track bugs and issues using JIRA.
PLE DGE, RAY, RIF, RuleML, iRods, RDF		There are a number of standard interchange formats for expressing repository policies that have been developed in recent months. The project will use these standards where possible and work with the MIT Digital Libraries to be informed in this area.

## 9. Technical Development

29. Technical development will follow SCRUM principals using open source techniques for collaboration and implementation. The majority of the source code will be Java although some scripting code may be used for support purposes. All source code will be hosted in a public SVN from the outset. The components will be part of a Continuum driven nightly build and outputs will be deployed to public beta servers as soon as they become available. In addition the project will use a Wiki for documentation and a JIRA tracker for issue tracking. All of this infrastructure is in production at CARET and will be made available to the project.

30. Adopting this rapid deploy early approach requires supporting infrastructure but improves visibility and keeps the development on track.

31. In addition the project will use online collaborative techniques to foster communication and reduce travel times.

## 10. Intellectual Property Rights

32. Software outputs of the project will be licensed under an Apache style license. Non-software outputs, including all the online project documentation will be licensed under Collective Commons 2.5 Share and Attribute. Copyright of the original material will rest with the originating institution of the individual, depending on the policies in force at those institutions. Where IPR is to be integrated with the Sakai or DSpace code base, that IPR will be contributed according to the agreed contribution policy of the recipient project.

## Project Resources

### 11. Project Partners

A letter of agreement was exchanged between Cambridge and UHI on 22 May 2007.

Partner	Role	Main Contact
CARET, University of Cambridge	Project Lead, VLE/VRE Developer Evaluation Researcher	John Norman
UHI	Fedora/VLE Developer Fedora Repository Manager	Sean Mehan
DSpace@Cambridge	Institutional Repository, DSpace Developer and Manager	Patricia Killiard
DSpace MIT	DSpace Consultant PLEDGE	MacKenzie Smith

### 12. Project Management

33. The project will be guided by a Steering Committee comprising representatives of CARET, the Cambridge University Library, UHI together with 1 external member from MIT. A place on the Steering Committee will be offered to the JISC Programme Manager. The Steering Committee will meet approximately every 8 weeks, as required. Executive responsibility for the project will reside with the Director of CARET, as Project Director.

34. An experienced project manager will perform day-to-day project management at 0.2 FTE and will report to the Project Director. In addition progress reports will be made using the progress report template of the JISC project management guidelines, it will be possible to track the project in near real time through the active publication of project outputs as they are assembled.

35. The project will be managed according to JISC project management guidelines, using the documents recommended in the guidelines. As part of the project closedown process, we anticipate that we will hold a debriefing workshop with those associated with the project, to capture any learning about the management and progress of such projects. It is envisaged that the project will have no training needs.

36. The project partners have existing project teams of experienced individuals from which the project will pull the most appropriate resources. The Lead representatives from each Organisation are listed below

Role	Individual	Organisation	Contact
Project Manager	Ian Boston	CARET	<a href="mailto:ian@caret.cam.ac.uk">ian@caret.cam.ac.uk</a>
Cambridge Libraries Lead	Patricia Killiard	University of Cambridge Library	<a href="mailto:johnf@caret.cam.ac.uk">johnf@caret.cam.ac.uk</a>
UHI Lead	Sean Mehan	UHI	<a href="mailto:sean@smo.uhi.ac.uk">sean@smo.uhi.ac.uk</a>
MIT Lead	MacKenzie Smith	Cambridge University Library	<a href="mailto:kenzie@MIT.EDU">kenzie@MIT.EDU</a>

### 13. Programme Support

37. The project will want assistance disseminating findings surrounding enhancement outcomes due to adoption of repository policies within the VRE/VLE environment. This might be in the form of supporting the Workshop or as opportunities to present at JISC events in this subject area. The project will also welcome the opportunity to collaborate with others working in this area especially other projects focusing on Workflow

### 14. Budget

There have been no changes to the budget from the agreed budget proposal. The Budget is attached in Appendix A.

## *Detailed Project Planning*

### 15. Workpackages

#### Phase 1

#### **Work Package 1 - Integration**

We will create an integration point between Sakai and DSpace and Sakai and Fedora that will allow read write access from the Resources Tool of Sakai into DSpace and Fedora. For DSpace, the integration will use DSpace LNI (webDAV) which will allow worksites within Sakai to deposit and reuse content from the Repository from within Sakai. A description of the work can be found at <https://camtools.caret.cam.ac.uk/access/wiki/site/jisc-ctrep/dspace/sakai%20integration.html> . We will also develop a similar integration component for Sakai and Fedora. The Integration will build on the ContentHostingHandler API in Sakai 2.4

Once the basic integration is performed we will work with the Team at MIT to make these components understand repository Policy as defined in RAY or RDF (an output of the PLEDGE project) so that simple deposit and reuse policies, defined by Work Package 2 can be communicated to the Sakai end used and enforced where appropriate.

We will work with the relevant Libraries and VLE/VRE teams to prepare the production installations of Sakai, DSpace and Fedora to accept the integration components.

#### Deliverables:

1. ContentHostingHandler Implementation with basic Policy support for DSpace/Sakai  
Responsibility: Caret
2. ContentHostingHandler Implementation with basic Policy support for Fedora/Sakai  
Responsibility: UHI
3. Preparation work for Production Deployment of DSpace/Sakai CHH  
Responsibility: Caret/Cambridge Library
4. Preparation work for Production Deployment of Fedora/Sakai CHH  
Responsibility: UHI

#### **Work Package 2 - Institutional Deployment and Policy Development**

We will work with Faculty and Department to identify communities of use and select/build suitable policies to address deposit and reuse within the Sakai/DSpace/Fedora environment. Having identified 1 community in each institution policies will be fine tuned to address the issues of that community down to the Item level. Policies that are developed will fall into the following 4 classes.

- Deposit Policies:
  - Inform Polcies: Policies that need to be informed to the end user, but do not require action

Project Acronym: CTREP

Version: 1.0

Contact: I. E Boston

Date: 18/06/2007

- Enforce Policies: Policies that represent some enforcement on the item being deposited at the point of deposit, eg Minimum Meta Data, Research Protocol.
- User Selected Policies: Policies that the user performing the deposit may select to be applied to the item at the point of deposit.
- User Supplied: Policies that are supplied by the user to express exceptional circumstances associated with the lifecycle of the item.
- Reuse Policies
  - Inform Policies: Policies that are attached to the item and communicated to the user on reuse.

The Deliverable in WP1 will attempt to provide rudimentary support for all these classes of policy, so that where policies are developed under this Work Package they may be made available to the selected communities.

Deliverables:

1. A community that is or will be using the local Sakai instance to work with the local institutional repository (1 for DSpace, and 1 for Fedora)  
Responsibility: Cambridge Library, UHI Libraries.
2. One or more policies to in the classes above to support those communities.  
Responsibility: Cambridge Libraries, UHI Libraries.
3. Deployed integration between Sakai & DSpace  
Responsibility: Caret, Cambridge Libraries
4. Deployed integration between Fedora & DSpace  
Responsibility: UHI

### **Work Package 3 - Evaluation**

We will work with the community identified in Work Package 2 to construct and evaluation framework to determine the impact of the deployment of WP1 and WP2.

Deliverables:

1. Report on impact of Phase 1 at Cambridge & UHI  
Responsibility: Caret, UHI

### **Phase 2**

Since the work in Phase 2 will largely depend on the outcomes of Phase 1 we cannot at this stage be more specific about the Work Packages in phase 2. The intention is use use Activity and Machine Workflow to address areas of inefficiency brought about by the increase deposit activity brought about by Phase 1

#### **Work Package 4**

Based on the findings in phase 1 identify a 3<sup>rd</sup> part activity based workflow implementation suitable for addressing phase 1 issues and integrate with DSpace and CamTools as necessary.

#### **Work Package 5**

Improve policies based on phase 1 and build workflows to automate the majority of the ingested items from CamTools.

#### **Work Package 6**

Using the metrics from WP3 build an evaluation to assess the impact of the workflow improvement and build a set of key metrics for future improvement cycles.

Appendix B contains the detail of the Work Plan

## **16. Evaluation Plan**

38. We will analyse a combination of individual logs, structured interviews, observation of work processes (and changes to processes) and usage statistics to assess the impact and suitability of the

Page 11 of 15

Document title: JISC Project Plan

Last updated: April 2007

DSpace/CamTools integration for the chosen study communities. A semi-quantitative research productivity model based on Wenger's Communities of Practice work will be developed and changes in activity recorded along the dimensions that determine productivity. This work will use qualitative analysis software such as ATLAS-T and NVivo. We expect to be able to demonstrate increased repository deposit activity and hope to be able to demonstrate increased productivity. A modest amount of time will be devoted to classic usability testing of the Deposit software with the chosen evaluation group.

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Phase1 WP3	Impact and effectiveness of the introduction of IR policies within the VLE/VRE	Detail to be determined in Phase1 WP3	Interview, Questionnaire, Observation followed by analysis as detailed above.	Increased Deposit activity with higher quality deposits
Phase2 WP6	Impact of improvements to Ingest and deposit workflows	Detail to be determined in Phase 2 WP6	Interview, Questionnaire, Observation followed by analysis as detailed above.	Reduction in workload in Ingest workflow, with more automated processing and fewer rejections or rework incidents

## 17. Quality Plan

39. The outputs of the project at Cambridge will become part of the existing production quality plan that manages the production environment for both CamTools and DSpace at Cambridge. As such the implementation will undergo continuous build and deployment, peer code review followed by a 6 week QA process resulting in 4 to 5 release candidates resulting in a final build ready for production deployment. Prior to that the implementation will use agile software lifecycle techniques with frequent releases and constant unit testing. We make use of nightly builds to deliver daily updates to prototypes and in build unit tests to ensure that code conforms to specification on every build.

## 18. Dissemination Plan

40. We will use open-source development practice of publishing (releasing) early and often, and we will support a project Wiki and Blog for general interest users and a JIRA site with anonymous SVN access for technical users. We will respond to feature requests and bug reports, as resources permit, but the site will allow community contributions in addition to our own efforts. We will attend and offer workshops at a minimum of 4 major meetings during the project, selected by agreement with the JISC Programme Manager. We will continue to contribute to both the Sakai Project and the DSpace project as part of the work undertaken in this project.

Timing	Dissemination Activity	Audience	Purpose	Key Message
Ongoing	Wiki/Blog	VLE/VRE Community, Libraries Community	To communicate the current status of work under the project.	Current status of the project, ongoing work and details of current outputs.
2 times per year	Presentations at community conferences	Communities involved with the supporting technologies, DSpace, Fedora, Sakai	To communicate the progress in the area, and get code adoption to support sustainability	

			longer term	
Once a year	JISC Libraries and Repositories Workshop	JISC Libraries and Repositories community	To communicate the outputs of the project and to encourage collaboration with other projects	Impact of automated ingest on research datasets and reuse. Impact of repository usage on reusable learning materials.
Once a year	JISC Workflow workshops	JISC Workflow community	To share experiences with implementation of Machine and Human workflows and supporting technologies.	

## 19. Exit and Sustainability Plans

41. CARET has committed to offering the Sakai platform to members of the University of Cambridge as part of its strategy for supporting teaching on the campus. We are also committed to supporting research groups at the University and have identified Sakai as the preferred tool. We believe strongly that the university sector can benefit greatly from shared software development based on the open-source model and have dedicated 2 FTE developers to supporting open-source projects. These developers have become valued participants in a number of open source projects including Sakai, being recognised by the award of one of only 2 Sakai Fellowships outside the US. Thus, the most important output of the project, - integration between CamTools and DSpace at Cambridge - will continue to be available for use, to research and education teams in all disciplines at the University. CARET will continue to develop and support the deployed software, and will champion its use.

42. CARET is an active member of both the Sakai Foundation and DSpace projects. All the outputs of the Sakai integration will be available as they are developed in the main Caret SVN repository under an Apache License. In addition the project will endeavour to ensure that the code and documentation is available as part of the core Sakai deployment or as part of the contributed modules, both in the Sakai Project SVN. Code entering the Sakai Project SVN will also be license under and Apache2 style license.

43. In addition the project will run both a Blog and a Wiki where all content will be made available as it is created under a Collective Commons license (Share and Attribute). A JIRA based issue tracker will be used to record and track issues within the project as they arise.

Longer term CARET will maintain the SVN, JIRA, Blog and Wiki as part of its continuing commitment to Open Source. When the community eventually moves on to other things, the outputs will be archived to DSpace@Cambridge

## Appendixes

### Appendix A. Project Budget

Directly Incurred Staff	March 07	April 07– March 08	April 08– March 09	TOTAL £
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]



Pledge based policy component. Output as an early Demo environment to support WP2 – Institutional Deployment and deployed in production to enable WP3 – Phase 1 Evaluation			
Work Package 2 Institutional Deployment	1/3/2007 – 18/10/2007	7 months	1 FTE Repository Manager (DSpace@Cambridge) 1 FTE Support and Liaison Officer (DSpace@Cambridge) 1 FTE Repository Manager (UHI) Consultancy support from MIT
Will identify and engage with communities requiring repository access. Will identify and specify licensing and policies for deployment. Will prepare for support of the communities in production.			
Work Package 3 Phase 1 Evaluation	1/3/2007 – 18/2/2008	11 months	0.25 FTE Researcher (CARET) 0.25 FTE Support and Liaison Officer (DSpace@Cambridge)
Engage with the target communities and build an evaluation framework that addresses the subject area of those communities. Once WP1 has delivered a production environment perform evaluation by interview, survey and observation. Analyse results using specified methodology and prepare Phase 1 review report for the Phase 1 review meeting.			
Phase 2	10/1/2008 – 3/12/2008	11 months	
Work Package 4 Workflow Integration	10/1/2008 – 30/9/2008	8 months	0.5 FTE Senior Developer (CARET) 1 FTE Developer (CARET) 0.5 FTE DSpace Developer (DSpace@Cambridge) 1 FTE Developer (UHI)
Based on the findings in phase 1 identify a 3 <sup>rd</sup> part activity based workflow implementation suitable for addressing phase 1 issues and integrate with DSpace and CamTools as necessary.			
Work Package 5 Policy and Workflows improvements	10/1/2008 - 18/10/2008	9 months	1 FTE Repository Manager (DSpace@Cambridge) 1 FTE Support and Liaison Officer (DSpace@Cambridge) 1 FTE Repository Manager (UHI) Consultancy support from MIT
Improve policies based on phase 1 and build workflows to automate the majority of the ingested items from CamTools.			
Work Package 6 Phase 2 Evaluation	10/1/2008 – 3/12/2008	11 months	0.25 FTE Researcher (CARET) 0.25 FTE Support and Liaison Officer (DSpace@Cambridge)
Using the metrics from WP3 build an evaluation to assess the impact of the workflow improvement and build a set of key metrics for future improvement cycles.			
Whole Project Lifecycle	1/3/2007 – 31/01/2009	23 months	
Work Package 7 Reporting	- 31/01/2009	23 months	0.1 FTE Project manager Support and contribution from work package teams
Manage the collection of outputs from WP1-6 for inclusion in the final report			
Work Package 8 Dissemination	10/1/2008 – 3/3/2008	2 months	1 FTE Support and Liaison Officer (DSpace@Cambridge) 0.25 FTE Researcher (CARET)
Organise and run a Workshop to be run after Phase 1 to disseminate findings to stakeholders and wider community.			