JISC DEVELOPMENT PROGRAMMES
Project Document Cover Sheet
Phase One Summative Evaluation Final Report

Project

Project Acronym	StORe	**Project ID**	
Project Title	Source-to-Output Repositories		
Start Date	1.9.2005	**End Date**	31.8.2007
Lead Institution	University of Edinburgh		
Project Director	John MacColl		
Project Manager & contact details	Graham Pryor graham.pryor@ed.ac.uk		
Partner Institutions	• University of Edinburgh (lead) • University of York, representing 'White Rose Partnership' • University of Birmingham • London School of Economics • University of Manchester • Imperial College London • University College London • UK Data Archive • Johns Hopkins University		
Project Web URL	http://jiscstore.jot.com/WikiHome		
Programme Name (and number)	Digital Repositories		
Programme Manager	Neil Jacobs		

Document

Document Title	Workshop Topic Guide		
Reporting Period			
Author(s) & project role	Meik Poschen (NCeSS, University of Manchester, Evaluator) Roger Slack (University of Wales, Bangor, Consultant)		
Date	10 September 2007	**Filename**	StORe_Ph1SummEvalReport
URL			
Access	✓ Project and JISC internal	General dissemination	

Document History

Version	Date	Comments
1.0	31/07/07	Draft
2.0	03/09/07	First revision (after presentation to John MacColl & Graham Pryor)
3.0	07/09/07	Second revision
4.0	10/09/07	Final revision

Phase One Final Report, WP 5: Evaluation of Pilot Demonstrator

Content

Appendix

1 Introduction

This document presents the final results of the phase one summative evaluation of the StORe[1] project's pilot demonstrator prototype system. As stated in the 'Summative Evaluation Plan v1.3' for WP 5 (see Appendix A), the "evaluation will assess the technical structure, functionality, design and quality of the demonstrator system, and the appropriateness of the 'common model' approach, using workshops to test the system with representative repository users".

In the recently completed two year StORe project a social science pilot demonstrator system has been built up and refined according to the stages of work (see http://jiscstore.jot.com/PilotMiddleware) at the UK Data Archive (UKDA). It federates a UKDA source repository and the Research Articles Online institutional output repository (ePrints) of the London School of Economics (LSE), based on the common/generic model, as outlined in the 'WP 3: Business Analysis' report. This business analysis report is one base document for the summative evaluation and itself based on data gathered in the 'WP 2: Survey of Researchers' (see http://jiscstore.jot.com/SurveyPhase).

The summative evaluation consists of a workshop topic guide and a review of the WP 3 business analysis report described in the chapters 2 and 3, both part of *deliverable one* already presented in two separate reports end of April 2007 (for more details on deliverables see Appendix A). It turned out to be very difficult to commit users to participate in the planned two workshops. In the end one workshop could be conducted with only three users taking part, who not even had been participants in the StORe project before. Due to these circumstances this final report is called phase one, as it is probable that a second complementary evaluation phase will be conducted in the future also based on the approach introduced here. This process and its repercussions are described in detail under 'Change Management' in chapter 4. Chapter 5 shows how the workshop session occurred eventually (part of *deliverable one*) and chapter 6 finally presents the results of the evaluation (*deliverables one and two*).

During the evaluation a close collaboration with the UK Data Archive in Essex was invaluable concerning the appropriate dates for the workshops, the technical prerequisites and demands for the user testing as well as respecting the knowledge and experience towards the pilot and its ongoing development and use.

2 Topic Guide

2.1 Objective of Workshop Topic Guide

Within the summative evaluation plan of WP 5 (see Appendix A), the topic guide is the basis of the workshop(s), and aims broadly at finding out barriers and facilitators to use and usability issues around StORe. The main contribution to this lies in the user testing of the social science pilot demonstrator system, followed by a questionnaire

[1] For an general overview of the JISC and CURL funded StORe project see http://jiscstore.jot.com/WikiHome

and a discussion session to derive users' themes and issues for the evaluation of the pilot in consideration of the common/generic model.

The workshop topic guide is part of *deliverable one* (together with organising and completing the planned two workshops and the review of the business analysis report) of 'WP 5: Evaluation of Pilot Demonstrator' within the StORe project. It presents the approach and underlying questions to conduct the user testing session(s) for the summative evaluation of the pilot demonstrator prototype system, the major outcome of the StORe project (executed in 'WP 4: Pilot Middleware'). The topic guide already has been presented as a separate deliverable end of April 2007 and was slightly adjusted for this final report.

2.2 Workshop Outline and Preparations

It was planned to carry out ideally two workshops, depending on the responses of the researchers invited. Each workshop should last approximately three hours with about eight users participating. The workshops should include (as already framed in the `Summative Evaluation Plan v1.2`):

- a presentation explaining project vision, concepts and objectives,
- the showing of a previously recorded video of the demonstrator system (< 10 minutes, recorded and explanatory speech added with the "Camtasia Studio 4" software) to explain its functionality and use,
- the user testing of this demonstrator system (which will be video recorded to provide retrievable data to inform reports and as a resource for developers, users and funders),
- the distribution of a (brief) questionnaire (Likert scale) to measure usability and user satisfaction of the pilot, and
- a discussion session to gain feedback on project vision, concepts and achievements.

Also it was planned for up to half of the users to attend not in person but via Access Grid (AG), if this would be more suitable. Participants should be selected by NCeSS from within the research groups at Essex (UKDA, with the help of Ken Miller) and London (LSE, with the support of Guy Burton) and preferably from attendees in the StORe questionnaire and interviews ('WP 2: Survey of Researchers'). The base of users at the Essex research group consisted of about 150 researchers, whereas the group of the LSE's researchers encompassed over 400 contacts made through the different survey phases from different institutions. A part of these users have also been involved in the ongoing user driven development of the pilot, i.e. having had the opportunity to get a login for the system and work with it. Therefore the invitations to attend the workshops were different for the Essex researchers (see Appendix B) and for the LSE group (see Appendix C), as the latter is combined with a reminder to make use of the pilot. The invitations have been sent out beginning with April 16 (for the further process of contacting potential participants for the workshops see chapter 4 'Change Management').

2.3 Initially Planned Workshop Agenda

It was initially planned to hold the workshops at the University of Manchester between the 8th and 18th May. In the following the proposed agenda and tentative schedule is presented (for the workshop agenda eventually put into practice see chapter 5):

10:00 – 10:15 *Arrival: Tea & Coffee*

10:15 – 10:20 Welcome and brief introduction round

10:20 – 10:30 Presentation explaining project vision, concepts and objectives

10:30 – 10:45 Probing of the users' so far experience with StORe or similar systems

10:45 – 11:00 Presentation of a short (< 10 min) video of the demonstrator system to explain its functionality and use

11:00 – 11:15 *Break: Tea & Coffee*

11:15 – 12:15 User testing of the demonstrator system (video recorded), followed by a query of initial reactions: "So what do you think?"

12:15 – 12:30 *Break: Tea & Coffee*

12:30 – 12:45 Distribution of Likert scale questionnaire with a brief introduction

12:45 – 13:45 Discussion session

2.4 Approach of the Topic Guide

The common/generic model (elaborated in the Business Analysis, WP 4 based on the findings of the survey, WP 2) is one main influence for the design of the pilot. For the evaluation it is important to identify categories of facilitators and barriers from the common/generic model as an underlying frame of analysis. A more detailed outline can be found in the review of the business analysis chapter (see chapter 3). Furthermore the robustness and sustainability of the system will have to be assessed after the workshops.

In keeping in mind the following topics, it is important for the evaluators to be careful not to overload the participants with things to discuss and not to imply the outcomes. Also these topics will be used for analysing the workshops afterwards, when sighting the video material and examining the memos.

2.5 Testing Session, Discussion and Feedback

2.5.1 Prior to the testing

The probing of the users' so far experience with StORe or similar systems prior to the testing should foster people getting in an appropriate state of mind for the issues of the pilot testing. Furthermore the evaluator is provided with additional information to better assess the outcomes of the testing session. Trying to covering the following areas should prove substantial to accomplish this:

- Current projects: probe for history, engagement with technologies and institutional background
- Types of collaboration involved in currently: probe especially for uses of repositories and experiences thereof (as both user and consumer)

- Perceptions of repositories in general: technical, organisational, professional and social issues – looking for barriers and facilitators to use.

2.5.2 Feedback directly following the testing

Initial comments: After the testing session it will be important to capture the initial reaction of the users to achieve feedback in form of direct genuine statements, expected or not. Therefore open questions are adequate, especially the start question: So what do you think?

2.5.3 Likert Scale Questionaire

The brief Likert scale questionnaire (see Appendix D) focuses on items to measure usability and user satisfaction of the pilot in a quantitative way to enhance the overall evaluation.

2.5.4 In the discussion session

Further topics to inquire on in the discussion session are:

- Asking for scenarios of use: how do people envisage using StORe in their projects and what barriers do they see (especially with technical, organisational, professional and social issues)
- Interim summaries given by the evaluator based on the users' scenarios will help to come to perceptions of the current state and possible developments of the pilot
- Closing remarks: by looking at the issues of the workshop so far, the evaluator elicits especially barriers, facilitators and any possible future use of the system

3 Review of the Business Analysis

3.1 Objective of the Review

The review of the business analysis ('WP 3: Business Analysis') is part of *deliverable one* (together with the workshop topic guide and organising and completing the planned two workshops) of WP 5. Like the topic guide, the review of the business analysis has been presented as a separate deliverable end of April 2007 and is included in this chapter in a refined version.

In analysing the findings of the business analysis (which is based on 'WP 2: Survey of Researchers') and under consideration of the appropriateness of the presented 'common model' approach, the review sets the theme for devising the workshop topic guide and the evaluation of the pilot demonstrator system as a whole. The review aims mainly at identifying facilitators and barriers of use for the development of the pilot demonstrator system. These are derived from the presented common/generic model (as after 'WP 2: Survey of Researchers'), which provided "the common ground between the disciplines" (Business Analysis, p.6), leading to user requirements for the pilot demonstrator system. Additionally the importance of the system's overall robustness and sustainability in use will be pointed out. To finally assess these issues the actual use of the system will be analysed in the workshop user testing sessions.

Three business analysis related documents are of further significance in this and will be taken into account when applicable:

a. the discipline variations document ('Discipline variations.doc' in WP 3), resulting from the survey phase and completing the business analysis
b. the technical specifications document ('TechSpec (4).doc' in WP 4), specifying technical requirements from the findings of the business analysis
c. the walkthrough document for the actual pilot demonstrator system ('SystemWalkthruV2.doc' in WP 4), introducing the system with screenshots and text description

The findings of this review incorporate directly into the workshop topic guide and, together with themes and issues unfolding from the workshops, will eventually contribute to the *third deliverable* of completing the demonstrator evaluation report and writing up a common/generic model evaluation report (see chapter 6).

3.2 Facilitators & Barriers in the Common/Generic Model

The basic assumptions from the common/generic model (business analysis and the discipline variations documents from WP 4, based on the findings of the survey from WP 2) lead to the functionality and design of the pilot according to the development stages of work in WP 4. For the evaluation of the pilot and particularly for the outline of the workshop topic guide it is important to identify categories of facilitators and barriers from the common/generic model as an underlying frame of analysis. Facilitators and barriers identified in the workshops can then also be compared with the list below, to find out what fosters the use of the system.

The added **comments** are derived from the discipline variations document and present possible constraints to the common/generic model. Including them here also presents a broader notion of which issues could appear in the workshops.

Category of facilitators

- Bi-directional links between source and output repositories are considered to be at least useful by the vast majority of the users/researchers.
  Comment: A minority of Astronomers were opposed to the aims of StORe, while two interviewees in the Biosciences even expressed hostility towards the project.
  Comment: Some researchers in the survey encountered problems with understanding the term 'repository', i.e. many of them in Chemistry; in the Social Sciences it seemed often that different kinds of repositories were indiscernible for the participants.
- The preferred method of searching was a simple search ("Google type").
- Self-reliance seems to be favoured by most researchers referring to data management and repository use.
- Common minimum metadata elements have been recognised as useful (e.g. author name, project title and description, keywords).
  Comment: In the Biosciences the term metadata seemed not to be widely familiar.

- Data-sharing is seen as a fundamental and important concept (especially between known individuals).

Category of barriers

- In many cases, institutional, library or other support was not acclaimed.
  Comment: Often participants of the survey did not know of such support.
- The coverage of repositories as well as standards regarding metadata, keywords and data formats was perceived as inconsistent.
  Comment: Discipline traits often seemed to promote such developments.
- "Each discipline identified barriers to the actual deposit of data/outputs, either because of time restraints, the bureaucracy imposed by repositories or constraints arising from their own or others' intellectual property rights." (StORe Business Analysis, WP 3, p.6)
- Data-sharing is seen as a fundamental and important concept, but more likely between known individuals than repositories.
  Comment: In Chemistry and Physics however, publisher repositories have been stated as most commonly accessed.

3.3 Robustness & Sustainability: Issues to explore

The concrete functionality of the StORe system, derived from the common/generic model, is concretised already in the business analysis and further specified in the technical specifications document (WP 4). In looking at this functionality on paper, the direction and course of the technical development is plausible. In furthermore including the walkthrough document (WP 4) into the examination, the screenshots and descriptions of the pilot in use follow the path outlined in the business analysis and the technical specifications.

Robustness and sustainability are two important criteria to assess the use of a system in terms of functionality, configuration and user's benefit. It comes down to the question, if users would use the system in the wild, for their everyday work. This means on the other hand, to find out the users' reasons of not wanting to use it. This also indicates robustness and sustainability are issues to explore. The workshops aim at finding out what users will say about using the StORe pilot demonstrator system.

4 Change Management

This section subsumes the change management which had to be implemented in the StORe summative evaluation because of difficulties in finding participants for the workshops. The initial three month evaluation was planned to take place between March and May 2007. The first wave of invitations (to over 500 addresses which participated in or have been connected to the StORe project before) to attend the workshops (see Appendix B and C; also see sub-chapter 2.2) was sent out as planned via email beginning with the 16th April 2007. With only one response within the two weeks following a second wave and even a third including an incentive scheme and the attempt to speak to people directly was implemented. The texts send out had been slightly changed in each wave and the emails had been sent out by different people and institutions, namely Ken Miller (UKDA), Guy Burton (LSE) and the evaluator

(NCeSS) himself. Within the third wave a second person responded, making up an extremely poor response of overall two users responding within six weeks to the different workshop invitations. These two people in the end could not even participate because they had not been available or interested anymore after some weeks.

Due to the problems in committing users to take part in the evaluation workshops the time frame was extended to end of June 2007 after reviewing the progress at the StORe Project Board meeting on 25th May 2007 in Edinburgh, leading to the third revision of the summative evaluation plan (see Appendix A; the second revision was implemented right at the start of the evaluation). But also with this extra time it was not possible to commit users from within the StORe project through the approach to "increasingly contact users at Essex, Manchester, Edinburgh and if feasible LSE through direct, personal contact, also using already existing contacts" as stated in the revised evaluation plan. This eventually resulted in acquiring only three participants in the end. The small number and time constraints rendered the fallback position of additional individual face-to-face sessions, as mentioned in the evaluation plan, useless. But it was at least possible to find one date for all users leading to a workshop on the 29th June 2007, which had to be slightly adapted to the new circumstances (see chapter 5). The final results and the first draft of this report have been presented to the StORe project director and project manager by the evaluator on 1st August 2007 and further been distributed by the project manager to the Project Management Group on 2nd August 2007. It was agreed to name the present report 'Phase One Summative Evaluation Final Report' with the intention to plan a second phase evaluation for the first half of 2008. Besides taking early measures to commit users from within the StORe project this additional evaluation shall make use of the same approach conducted by the same evaluator in order to complement the hitherto findings presented here. Important issues described in the topic guide and review of the business analysis (see chapter 2 and 3) therefore could not elicited to its full intend. These inevitable shortcomings in assessing the common/generic model due to the lack of committed StORe participants in the phase one evaluation are illustrated further in chapter 6 ('Evaluation Results'). The second phase aims at resolving these issues properly.

5 The Workshop

The workshop took place on Friday 29th June 2007 between 13:30 and 15:30 in the Access Grid Node 1.10, Kilburn Building at the University of Manchester. From the three participants two attended via the Access Grid from a node at the UK Data Archive, University of Essex organised by Ken Miller and Mike King from the UKDA, while the third user flew in from Edinburgh and took part in the Manchester AG node together with the evaluator.

The workshop consisted of the following, compared with the initial slots slightly adapted agenda (cf. sub-chapter 2.3), whereas all times depict the effective time slots on the day:

13:15 – 13:40 Set up of the AG node, connection with Essex and welcome to the Edinburgh participant

13:40 – 13:50 Official welcome and brief introduction round including probing of the users' so far experience with StORe or similar systems

13:50 – 14:05 Live presentation of the StORe system

14:05 – 15:10 Pilot user testing session

15:10 – 15:30 Feedback session [and group discussion]

The Likert questionnaire was exchanged via email after the workshop in a clickable version. An additional change was the decision not to use a recorded video for the presentation of the StORe system, but a live presentation done by the evaluator himself. The reason was mainly to be more flexible to adapt the presentation to the knowledge of the participants who – other than initially planned – had not been familiar with the StORe system in any way so far. Also, especially given the shorter time and the smaller group size it seemed more appropriate to the evaluator to do the presentation in a more direct and personal way in order to try to establish a basis for a lively interaction between evaluator and participants.

The Manchester Access Grid system connection with the Essex AG node crashed unexpectedly for unknown reasons right before the evaluator planned to end the testing session at around 15:10 (about 60 minutes had been scheduled for this slot, after 65 minutes the participants were still motivated to explore). The crash did not affect the recording, which went on with the black Essex screens but still recording the Manchester site. For that reason the two participants in the Essex node had not been able to take part in the initial feedback session at the end of the workshop and also a group discussion was not possible. But this was the only harm done as introduction and testing session worked as planned, or as Ken Miller put it in an email afterwards: “AGN ended a bit abruptly, if you said goodbye, sorry we didn't respond, but we didn't hear. We were obviously too busy evaluating StORe.” Feedback from the two Essex users was later obtained via telephone interviews (see next chapter).

6 Evaluation Results

This chapter describes the results of the summative evaluation namely *deliverable two* (the compiling of themes and issues unfolding from the workshop) and *deliverable three* (the assessing of the common/generic model) of WP 5. Both deliverables are for the first time presented in this report and complete the phase one evaluation. The composition in the following furthermore draws on the approach depicted in the workshop topic guide (see chapter 2).

6.1 Before the testing session

After a brief introduction round and a question about the users’ so far experience with StORe or similar systems the users agreed in being recorded in the workshop.

Participants

- The user from Edinburgh works as a Data Librarian at the University and also in a JISC project (“DISC-UK DataShare”) in the repositories and preservation programme. He explained to know about the StORe project, but not about the prototype and to be “interested in its possibilities”.
- The two Essex users come from a Social and Economic research institute and have not worked with similar systems before. When working with data sets it is mostly

about scientific quantitative data. They stated that they were "curious" and interested to "find out what the system can do".

The evaluator gave a live presentation of the system partially on basis of the system walkthrough document ('SystemWalkthruV2.doc' in WP 4) to introduce the pilot to the users (also see chapter 5). No participant had any questions before the testing session. Each user is handed a workshop tasks document (see Appendix F) to guide him through the testing session. The evaluator emphasised that this document should be just a guideline and not intended to keep the users from exploring the system by too strictly following it.

6.2 Communication during the testing session

Due to the small group size, i.e. evaluator and one user in the Manchester AG node and developer and two users in the Essex node it was possible for the users to ask questions and make remarks and to get answers and help from the developer and the evaluator. For the evaluation this proved to be very fruitful, as can be seen in the following topics. The information provided here from the interaction in Manchester is more detailed because it was more frequent and due to the direct participation and note taking of the evaluator present. The interaction with the Essex site were recorded on paper after the evaluator reviewed the recording.

Manchester topics

- Concerning metadata it became obvious, that categories had to be clearer to point out what is meant exactly by source, contributor, coverage etc.
- Furthermore the interlinking between the pages, i.e. the navigation concerning the metadata page should be better (e.g. "how to go back from here to the start page?").
- The representation of collections and folders and the underlying concept of collection, item and folder was unclear: a collection is a folder, but then again it is possible to add a folder (to structure) within a collection, which was felt to be "a bit confusing". Additionally, creating a folder in a collection requires the metadata again, which confused the user again, because he did not know, if he had to add this and if he had done so, would it have to be the same as the collection's parent metadata?
- About metadata: "What does custom metadata mean exactly?"
- The ingest function was not immediately understood.
- The mechanism of publishing something with the command move seemed unclear, and maybe "this could be represented as a main command on the first level" (where the main tabs are).
- The concept of moving, i.e. publishing an item was not apparently clear: who will approve and verify in what case and what collection (private, public, collaborative).
- The question of access rights in collaborative collections emerged: what will happen to a file, if a right is revoked from a collection, i.e. folder?
- It was the observation, that after a while and in the real working state the list of collections will get longer and it would be nice to have something to keep an overview and keep the list passably slim, maybe with categories.

- Embargo was seen as a very good and interesting function. But some questions remained: "When will it be approved, before or after the embargo time" – whereas it was the opinion that "after" would make the most sense.

Essex topics (user 1)

- In answer to the question "Could I make private collections public?" the developer explained something about private/public/collaborative collections (with the details inaudible) and answered: "No."
- The question about adding items (files) "Which files can I add?" is answered with "any".
- The user had a slight problem with using a LSE study number (the only user to try a LSE study number in the session), after some help from the developer it worked and the user clicked on the SN and viewed the content with a satisfied "Ahh!"

Essex topics (user 2)

- The developer explained the question "Why do people create the data set/a new data set?" resulting from an unfamiliarity with the overall concept of StORe.
- Entering the type of research and the study number in the metadata fields made problems (user asked and developer helped).
- The developer explained the ingest functionality after being asked.
- The user asked how files could be approved as he could not find one file from a collaborative collection to find out that files can be approved by himself, the user.

6.3 Feedback session

As described in the workshop topic guide it was planned to elicit initial comments from the participants to get first unfiltered statements right after the testing session (see sub-chapter 2.5). After the crash of the Access Grid connection with Essex this was only possible with the participant on-site in Manchester. Therefore naturally a discussion session could also not take part and the evaluator probed for the relevant topics like use, barriers and facilitators. With the two Essex users this data was raised through telephone feedback sessions as described in the next sub-chapter.

Manchester user

- What do you think?
  → "Generally it is quite self-explanatory, given the task and this testing session."
  → "Explanatory tags would be good to give context", i.e. a help function.
- Could you envision using this system?
  → For his own work the user qualified the system as useful and added that he would be "not a researcher" but that from a "project perspective" StORe could fit into the portfolio, maybe in a modified (more mature) version for collaboration.
- Barriers/facilitators?
  → He assessed it "difficult for someone coming to this new", without the background from participating in the project or as a researcher, therefore for him it was "difficult to envision it as a functioning resource in this test". He added that the testing session naturally would be a bit out of a "real" context, but that found the demonstrator to be "very useful".

Being the only two persons physically present in the Manchester AG node, user and evaluator also exchanged remarks, questions and answers during the testing session (see sub-chapter 6.2). This forestalled a lot of important information which otherwise would have had to be probed in this feedback session.

6.4 Telephone feedback sessions

In order to get the necessary feedback telephone interviews have been conducted as prompt as possible a few days after the workshop with the participants from Essex (the workshop took place on a Friday and in its aftermath it was logistically not possible to do this on the same day). Here the same issues were probed as in the initial feedback session (What do you Think? use, barriers/facilitators).

Essex (user 1)

- What do you think?
  → "I enjoyed the demonstration and generally am impressed with the potential that this offers."
- Could you envision using this system?
  → It is a "good tool, could imagine using it." The user states that he so far has not worked with similar tools.
- Barriers/facilitators?
  → "It's cumbersome sometimes with all the folders and subfolders."
  → Also it was said that it would be good for a private collection (originally the participant used the word "project" instead of collection) to have a function to open it up to public later.
  → On the other hand it was felt that the system is easy to use and navigate.

Essex (user 2)

- What do you think?
  → The system was seen as a "very interesting tool".
  → The testing itself was assessed as "ok", but because of the "not real" testing session ("a real example would be good") sometimes not transparent and difficult to see "what was going on".
- Could you envision using this system?
  → No concrete use was seen by this user, but with its collections (also the second Essex participant originally used the word "project" instead of collection) the tool should be "nice for people who feel the need to collaborate" and who are therefore working with datasets which they want to share.
- Barriers/facilitators?
  → As with the other Essex user also in this case the wish arose to be able to make a private collection public, i.e. to change such properties of collections after having them created.
  → The concept of being able to add subject tags (in the metadata information page) which also possibly erases the old subject(s) and the way of adding more than one tag was felt to be complicated and with no consistency in classification ("it's a mess").
  → The metadata entry page as a whole was criticised for having not quite clear categories – the core problem in that was seen in the missing help function to explain the fields.

→ As the command to view or edit collection information was placed on the right side of the screen but the rest of commands on the left side this user found navigating and keeping track difficult in this instance.

6.5 Likert Scale Questionaire

The brief Likert scale questionnaire – in a clickable version – was emailed to the participants immediately after the workshop and sent back by all three shortly thereafter. The overall results reflect the impressions of the testing and feedback sessions (see appendix E). The average also noted there is statistically seen a merely academic value, as the size of the sample is only three. But apart from the items B and F it shows a tendency which makes sense because all answers are in the area of 2 (*agree*) and 3 (*undecided*) and the users' frequencies here are homogenous:

- A. The menus and buttons **are** easy for me to understand *(two agree)*
- C. The tool **is** easy and simple to use *(two agree)*
- D. The results I obtain **are *in some degree*** clear and understandable *(two undecided)*
- E. The tool's options **are** easy for me to understand *(two agree)*
- G. I think the tool **will be** useful to me in my work *(all agree)*

More complex are the answers to the items F and B.

- F. The error messages **are *absolutely*** clear
  vs.
  F. **Not sure, if** The error messages are clear
  (one strongly agrees and two are undecided)

Also taking into account the evaluator's observations of the pilot testing it can be said, that at least some error messages could be clearer (see next sub-chapter).

- B. I understand what the pilot system will do ***!?***
  (one strongly agrees, one agrees and one disagrees)

B pictures the most heterogeneous results, with two answers positive and very positive and one negative. At the same time this is the item with the highest demands regarding the user's general understanding of the whole StORe context. With – other than planned – all three users not being involved in the StORe project in any way and two of them not even coming from a declared repository/library angle it is difficult to say, why the answers are like this: Because of missing context knowledge or due to the use of the system itself. Therefore this question is methodologically not valid in the sense, that it does not unambiguously measure what it ought to do, which should have been the understanding of the system in use.

6.6 Common issues arising from the testing session, feedback sessions and questionnaire

As the user base in this phase one evaluation was very small with three users, the issues arising from the workshop as a whole (collating testing session, feedback sessions and questionnaire) can not be easily generalised. A need for further prioritisation therefore becomes immanent and is planned to be achieved with a larger user base complementing this data in the second evaluation phase. Nevertheless it can

be said that the data surveyed through the different methods within this evaluation fit together very well and immanently shows no inconsistencies, but clearly indicate tendencies of how to improve the StORe system in a next step. The following issues are ranked roughly in the order of importance, which here means combining the quantitative number of incidents and its perceived qualitative weight for hindering the proper use of the StORe system.

- Collaboration is a very important asset of the StORe system. The test session was appreciated as very interesting and even joyful in the second half (especially after discovering the collaboration aspects). Albeit it was noted explicitly by two of the three users that the testing session naturally could not represent a "real" working situation and that it would be good to work with and try out the system in such.
- Metadata page 1: categories (fields) were felt to be unclear (not everyone is familiar with Dublin Core), especially here a help function is recommendable, e.g. giving examples of fields to fill in when moving over them with the mouse pointer. Especially "research type" and "study number" need some help material as the most errors have been triggered here and even the same error repeatedly.
- Metadata page 2: making changes opens the metadata page again with a message. This often left the users a bit perplexed as they expected a different page, e.g. the collection information overview. Especially in that case the navigation/interlinking between the pages could be clearer and also better implemented to solve this issue.
- The concept and terminology of moving (= publishing) an item and then approving/verifying it was not clear cut for the users.
- Error messages could be clearer and more visible. Despite of the error message and the marked fields on the metadata page it took two of the three users two to four tries to make the changes because of omitting one of the denoted fields, mostly "research type" and "study number". Also the only bug in the session was found: When trying to add a URL omitting "http://" a complicated error message was shown, the item was not added but the metadata page instead opened anyway; on occasions users did not realise this and wondered where the supposedly added file would have gone.
- A help function in general is desired and needed by all users. This could be a contextual help function via appearing tips when the mouse pointer is placed over an item and/or a help function realised as a link which is always visible and opens a page or additional window with explaining text. It showed to be important to have help especially for:
  - the different collections
  - research type and study number
  - Dublin Core metadata fields
  - custom metadata
  - ingest
  - approval queue (moving, approve, verify)
  - access rights
  - embargo
- The names and concept of collection, folder and adding items could be clearer. As a collection is represented by a folder and at the same time a subfolder of a

collection also is a real folder this tended to confuse every user (collection was more understood as a project and often also called so). Maybe a change of the collection symbol would solve this, additionally a function explaining something of the general concept of the system would be helpful. Furthermore a file is an item whereas a URL also is an item, the names of the commands (adding an item) and the following choice of adding a URL or a file at the same place was difficult to comprehend at first.

- The custom metadata function was not clear to any user.
- The ingest function was very much appreciated but took a while to see through, the name was not outright clear.
- The navigation was appreciated as easy usable, except with the metadata page and with the links getting onto the metadata page.
- It was not easy understandable how access rights exactly work.
- The sense and use of the embargo function should be made clearer.

6.7 Assessment of the common/generic model

Shortcomings of the approach due to the lack of committed users

As mentioned before in this report (see chapter 4): Especially for the assessment of the common/generic model presented in the business analysis document of WP 4 together with the in the document's review identified possible facilitators and barriers (see chapter 3) the need for more exploration in phase two with appropriate buy-in from end user partners is apparent. It will be necessary to assess the common/generic model from the perspective of the "generic" and involved StORe user, which could not take part in this evaluation's testing.

Despite these shortcomings it was possible to start to evaluate other important criteria for "the use of a system in terms of functionality, configuration and user's benefit" (see 'Robustness & Sustainability: Issues to explore' in sub-chapter 3.2) and show important barriers and facilitators in this respect in the last sub-chapter.

Robustness and sustainability of the system

The system proved to be robust (just one bug appeared during the testing), did exactly what was promised and the users could work well with the system given the not "real" situation (cf. especially feedback sessions in 6.3 and 6.4). Despite the list of things to improve (see last sub-chapter), the questionnaire shows a general satisfaction with the use and usability of the StORe system, which is corroborated through the observations of the testing session and the users' comments in the feedback session.

In face of the methodical problem of the users' not having been involved in the StORe project until the workshops (and with the too small number of participants to make the results scientifically significant) the system proved to be sustainable in answering the question, if users would use the system in the future for their everyday work. All users could imagine using the StORe system especially in the context of collaboration in a future, more mature version.

Appendix A – Summative Evaluation Plan

JISC DEVELOPMENT PROGRAMMES
Project Document Cover Sheet
Summative Evaluation Plan

Project

Project Acronym	StORe	**Project ID**	
Project Title	Source-to-Output Repositories		
Start Date	1.9.2005	**End Date**	31.8.2007
Lead Institution	University of Edinburgh		
Project Director	John MacColl		
Project Manager & contact details	Graham Pryor graham.pryor@ed.ac.uk		
Partner Institutions	• University of Edinburgh (lead) • University of York, representing 'White Rose Partnership' • University of Birmingham • London School of Economics • University of Manchester • Imperial College London • University College London • UK Data Archive • Johns Hopkins University		
Project Web URL	http://jiscstore.jot.com/WikiHome		
Programme Name (and number)	Digital Repositories		
Programme Manager	Neil Jacobs		

Document

Document Title	Summative Evaluation Plan		
Reporting Period			
Author(s) & project role	Rob Procter (Project Board, NCeSS, Univ. of Manchester), Meik Poschen (NCeSS, University of Manchester), Roger Slack (Univ. of Wales, Bangor, partner of NCeSS)		
Date	30 May 2007	**Filename**	StOReSummEvaluationPlan-NCeSS
URL			
Access	✓ Project and JISC internal	General dissemination	

Document History

Version	Date	Comments
1.1	Nov 2006	First Revision
1.2	19/03/2007	Second Revision (after Project Management Group meeting)
1.3	30/05/2007	Third Revision (after Project Board meeting)

Summative Evaluation Plan v1.3

Purpose

To evaluate the demonstrator system and project 'common model' in order to make recommendations for future development work.

The evaluation will assess the technical structure, functionality, design and quality of the demonstrator system, and the appropriateness of the 'common model' approach, using workshops to test the system with representative repository users.

Methodology and change management

Workshops: The researcher (Meik Poschen, NCeSS) will convene one to two workshops (or, as a fallback option, individual one-to-one testing sessions and interviews, also see below) and compile and use a topic guide (to be agreed with management committee) concentrating broadly on barriers and facilitators to use and usability issues around StORe. The first final versions of the topic guide and review of business analysis deliverables were delivered on schedule due to end of April 2007 as stated in this document.

Each workshop will last approximately three hours and involve about eight users. It was planned to select users from within the research groups at Essex (UKDA) and London (LSE) and preferably from participants in the survey of researcher use of repositories. Due to the extremely poor response (two users responded within six weeks) to the workshop invitations despite using different approaches including an incentive scheme, the procedure has to be adjusted to the new situation, as decided at the Project Board meeting on 25 May 2007. It is now planned to increasingly contact users at Essex, Manchester, Edinburgh and if feasible LSE through direct, personal contact, also using already existing contacts where applicable. The option of having the users not resident in Manchester participate via Access Grid, especially through an AG node at Essex currently is explored.

As already mentioned above, a fallback option would be, to conduct one-to-one testing sessions and interviews in case there still will be not enough participants for the workshops. The researcher then would visit users individually and proceed according to the workshop guide, which will keep its topics and agenda, but be slightly adjusted to the face-to-face sessions. The completion date of the evaluation will be extended one month until the end of June 2007.

The workshops (or individual testing sessions) will be structured as follows:

a. Introduction, presentation explaining project vision, concepts and objectives;
b. Presentation of a short (10 minutes<) video of demonstrator system to explain its functionality and use;
c. User testing of the demonstrator system. This will be video recorded in order to provide retrievable data to inform reports and as a resource for developers, users and funders; the testing will be bracketed by a probing of the users experience before and a query of initial user reactions afterwards;
d. Distribution of Likert scale questionnaire to users to measure demonstration system usability and user satisfaction;
e. Discussion session (led by international consultant) to gain feedback on project vision, concepts and achievements.

Deliverables

1. Workshops: user testing to identify usability issues, barriers and facilitators to adoption of common model.
2. Demonstrator evaluation report summarising user test results and issues arising from discussion session.
3. Evaluation report on acceptability of the 'common model' approach.

Timetable

March to June 2007

Note: At first three months, March to May 2007, now extended one month.

Month and *deliverables* in detail:

1. Generate topic guide; organise workshops with partner sites; make video of demonstrator system in use; review business analysis report;
2. Complete topic guide and review of the business analysis (*deliverable one, parts 1 and 2*); ongoing workshop organisation (especially finding participants);
3. Change management: ongoing workshop preparations (especially finding participants) and creating of fallback option (individual sessions);

Finish workshop preparations, conduct workshops (*deliverable one, part 3*) and commence writing up of document on themes and issues arising from workshops (*deliverable two*); complete demonstrator evaluation report and write up 'common model' evaluation report (*deliverable three*);

Appendix B – Workshop Invitation UKDA Users

StORe pilot user testing workshops

Dear Madam/Sir,

As a participant of the StORe user survey you already know about the Source-to-Output Repositories project (StORe) and its aims (for more information see http://jiscstore.jot.com/WikiHome).

Your contributions have been invaluable for the current phase of the project, the user driven development of the pilot demonstrator system, a portal which connects source repositories at the UK Data Archive with the Research Articles Online institutional output repository (EPrints) of the London School of Economics (LSE).

In order to evaluate the prototype with your further input, the National Centre for e-Social Science will conduct two pilot user testing workshops at the University of Manchester (your travel expenses can be reimbursed). We will need you participation to develop the system for your needs.

The two workshops each require about eight participants and will take approximately three and a half hours. After a brief introduction to the project and the workshop you will be shown a short video of the functionality of the StORe pilot system. Then you will have the opportunity to test the system, which will be video recorded for further evaluation of the system in use. We will treat all this information strictly confidential. Finally all users can exchange their opinions of the testing session, system's benefits, wishes and things to improve in an open table discussion. And last but not least coffee, tea, biscuits and cake will be provided for everyone's well-being.

The workshop can also be attended via Access Grid by about half the participants of each workshop, which also will be recorded.

It is planned to hold the workshops between the 8th and 18th of May, if you are interested we would set up an electronic calendar for you to enter you preferred dates.

We are looking forward to hearing from you, please send a reply of this mail to Meik.Poschen@manchester.ac.uk if you are interested and/or have additional questions and we will come back to you shortly.

Many thanks and kind regards

Meik Poschen

National Centre for e-Social Science (NCeSS)
University of Manchester
Phone: +44 161 275 1384
Email: meik.poschen@manchester.ac.uk
Web: http://www.ncess.ac.uk/about/people/Meik

Appendix C – Workshop Invitation LSE Users

Dear Colleague

I am writing to you regarding the Project StORe pilot IT project that has been designed to improve researchers' use and interaction of source data and research.

First, I would like to thank those that have so far taken part in making use of the IT pilot system and would encourage those of you who have not yet done so to sign up. The system forms the final phase of the cross-disciplinary and non-commercial Project StORe (Source-to-Output Repositories) and is now available online at http://oai.esds.ac.uk:8080/elated/. The system has been designed to encourage researchers to deposit and access data and material relevant to their research interests. Individuals interested in taking part should contact Ken Miller of the UK Data Archive at millk@essex.ac.uk in order to receive a username and password.

In addition, I would encourage you to provide feedback with regard to the IT pilot system. This may be supplied in one of two ways: either by opening up the link at the bottom of each page on the demonstrator or via the StORe Project's Wiki at http://jiscstore.jot.com/WikiHome.

Second, I would like to invite interested researchers to participate in the evaluation of the IT system through two pilot user workshops at the University of Manchester (your travel expenses would be reimbursed). The workshops, scheduled between 8-18 May, each require about eight participants and will take approximately three and a half hours.

The workshops would involve a brief introduction to the project and the workshop and a short video of the functionality of the StORe pilot system. Researchers would then have the opportunity to test the system, which will be video recorded for further evaluation of the system in use. We will treat all this information strictly confidential. Finally all users can exchange their opinions of the testing session, system's benefits, wishes and things to improve in an open table discussion. Individuals interested in participating in one of these workshops should contact the technical evaluator, Meik Poschen at the National Centre for e-Social Science (NCeSS), University of Manchester at Meik.Poschen@manchester.ac.uk.

Thank you for your assistance in this matter. Should you have any queries about the pilot system or Project STORE in general, you may either visit the project's website below or contact me.

Yours sincerely

Guy Burton

STORE Project Officer
Information Services
Library
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7955 6909
Email: G.J.Burton@lse.ac.uk
Website: http://jiscstore.jot.com/WikiHome

Appendix D – Workshop Likert Scale Questionnaire

Pilot User Testing Workshop Questionnaire

Many thanks for participating in the StORe user testing workshop so far. After using the pilot in the testing session before, we would like to ask you to fill out this brief questionnaire.

Please indicate the extent of your agreement or disagreement for each of the statements below by placing a tick in the appropriate box, according to the following scale:

1 Strongly Agree – 2 Agree – 3 Undecided – 4 Disagree – 5 Strongly Disagree

A. **The menus and buttons are easy for me to understand**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

B. **I understand what the pilot system will do**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

C. **The tool is easy and simple to use**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

D. **The results I obtain are clear and understandable**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

E. **The tool's options are easy for me to understand**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

F. **The error messages are clear**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

G. **I think the tool will be useful to me in my work**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	☐	☐	☐	☐	☐	

Thank You!

Appendix E – Results of the Clickable Workshop Likert Scale Questionnaire

Pilot User Testing Workshop Questionnaire

Many thanks for participating in the StORe user testing workshop so far. After using the pilot in the testing session before, we would like to ask you to fill out this brief questionnaire.

Please indicate the extent of your agreement or disagreement for each of the statements below by placing a tick in the appropriate box, according to the following scale:

1 Strongly Agree – 2 Agree – 3 Undecided – 4 Disagree – 5 Strongly Disagree

A. **The menus and buttons are easy for me to understand**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	3, 2, 2; Ø 2.3					

B. **I understand what the pilot system will do**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	2, 1, 4; Ø 2.3 !?					

C. **The tool is easy and simple to use**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	2, 3, 2; Ø 2.3					

D. **The results I obtain are clear and understandable**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	3, 2, 3; Ø 2.6					

E. **The tool's options are easy for me to understand**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	3, 2, 2; Ø 2.3					

F. **The error messages are clear**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	1, 3, 3; Ø 2.3 !?					

G. **I think the tool will be useful to me in my work**

Strongly Agree	**1**	**2**	**3**	**4**	**5**	**Strongly Disagree**
	2, 2, 2; Ø 2					

Thank You!

Appendix F – Workshop Tasks

Pilot User Testing Workshop Tasks

Thank you for participating in the StORe user testing workshop. Below you find tasks to guide you through the testing session. Please work on each task one after another. Take your time and feel free to explore.

➔START: StORe Home Page http://oai.esds.ac.uk:8080/store/

1) Browse through the collections and view collection information to see the metadata

2) Explore the Search and Advanced Search to find a document (enter metadata you viewed before) or just try “test” as a keyword

3) Login (case sensitive), change you password and enter your or fictional user data (name, surname, email)

4) Make yourself familiar with the functionality, i.e. the different tabs and commands and explore, what can be done

5) View your collections and create a new collection (private or public or collaborative) → If your type of research is based on existing data, please do not use the IPCSR Michigan repository. You can use e.g. the following as valid study numbers: UKDA => **100** / LSE => **00000750**

6) Edit and add metadata to your collection

7) Add a folder and an item to your collection (if you want to add a file, you can create an empty file on your desktop with ‘right-click->new’ and choose this)

8) Edit the item, add some metadata

9) Move the item for publication

10) View the approval queue

11) Freestyle: try out and explore what comes into your mind