



## JISC Project Plan Template

*The Project Management Guidelines have detailed instructions for preparing project plans.*

*Expand tables as appropriate.*

*Fill in the information for the header, e.g. project acronym, version, and date.*

*Prepare a cover sheet using the cover sheet template and attach to the project plan.*

### Overview of Project

#### 1. Background

*Summarise the background to the project (and how it builds on previous work) and the need for it (and why it's important).*

This project will explore the feasibility of developing data exchange models and data conversion tools for primary research data collected in the course of empirical research. Data conversion and proprietary data entry and analysis are particularly important and problematic aspects of data management and curation. Much important primary research data is created every day in the course of academic and policy research. While Data Sharing Policies are encouraging sharing and formalised archiving of data, the ideal life cycle for data creation to re-use remains beset by obstacles. The main issues involve the buying-in to a dedicated analytic strategy and typically a particular software package. Over the years the UKDA has seen a number of such softwares quickly become obsolete. To address the problem of incompatibility between software various data conversion tools have come of the market. However in the qualitative data analysis software field there are no such inter-software conversion tools. Open data exchange formats are necessary for maximising the opportunities for data sharing and long-term archiving.

This project aims to provide researchers and support staff working with primary research data with a suite of tools that will enable data to be long-term curated and exchangeable. The tools will research and both develop and test tools for contemporary qualitative data typically used by social researchers. The data formats to be included are those that are commonly used in qualitative research eg word, rtf and the CAQDAS package formats eg Atlas-ti, MaxQDA and Nvivo. The test data selected for this project are from the social sciences, but these formats are typically found across all domains of primary research.

The work to be undertaken in this project build on work already undertake by the investigator and other collaborating teams on two new and roughly specified models that have been developed recently. A draft but limited formal definition of a common XML vocabulary and Document Type Definition (DTD) based on the Text Encoding Initiative (TEI) for describing these structures has been prepared by ESDS Qualidata<sup>1</sup>. The Universities of Melbourne and Queensland have developed a draft Qualitative Data Interchange Format (QDIF) for e-Social Science (QDIF). Both centres have been working closely together in this very early development phase, but as yet, neither have any dedicated funding to work further on realistic development or testing, so the specification work remains on the back-burner.

The model is aiming at specifying in XML or RDF, a data exchange format that can represent a data collections and retain links between data, annotations and related objects. The model makes use of the DDI and TEI for metadata, and is compliant with Dublin Core and OAI. The model will be tested on various data types including output to a generic archival format and export from (and possibly import to) 2-3 of the brand leaders of the CAQDAS software. Some basic import/export tools will be developed. Agreed testers of the models ad tools are

A small scale evaluation of the models and tools will be carried out in order to inform JISC of the most viable options for future development in this area. A longer-term aim of this work would be to build a fully functional and scalable facility or service where data formats can be submitted and seamlessly returned in a chosen, desired format. The work in this proposal aims to lay the foundations upon which such a sustainable service could be built.

## 2. Aims and Objectives

*List the broad aim or purpose of the project, and the specific objectives you intend to achieve.*

This project aims to provide researchers and support staff with standards and associated tools for data (and metadata) format conversion. Central to these aims are the recognition and use of best practice in longer-term data management and curation.

The specific objectives are to:

- research and develop a numeric data exchange standard and conversion tools;
- research and develop a qualitative data exchange standard and conversion tools;
- test and evaluate these standards and tools;
- assess the feasibility of developing a web-based service for data conversion based on these tools sets.

## 3. Overall Approach

*Describe the overall approach you will take to achieve the objectives outlined above, including:*

- *Strategy and/or methodology and how the work will be structured*
- *Important issues to be addressed, e.g. interoperability*
- *Scope and boundaries of the work, including any issues that will not be covered.*
- *Critical success factors.*

There is a PI and one full time officer working on the project. The main work will be compiling and assessing all relevant schema, key softwares and data structures, and finally complex datasets.

The work is limited to:

1. assessing all available relevant schemas
2. consulting with a range of key CAQDAS software suppliers to look at willingness to buy-in and the feasibility of developing import and export tools
3. a proposed draft exchange format
4. assessment and evaluation of the draft exchange format by UKDA and agreed collaborators on selected datasets
5. developing some limited export tools from CAQDAS software to the proposed format
6. workshop with suppliers and data archivists/curators to assess potential level of interest in standard
7. assessment of whether a web-based data conversion service might be possible

On the last point, once these demonstrator tools can be shown to work it should draw attention to the potential for a fully specified statistical data curation/exchange web service. Users could upload data that had been extracted from a software package and return a chosen converted format. This is a logical extension to the work that would be funded within this bid and is would be highly beneficial low investment tool for institutional repositories. The short time scale and limited budget for this proposal does not allow for full exploration of this kind of facility, but the potential will be discussed in the final report.

## 4. Project Outputs

*List the tangible deliverables (including reports) your project will create, and the less tangible knowledge and experience you hope to build and share.*

The main deliverable will be a model for a data exchange format for qualitative data – focussed on text and a limited set of import/export tools.

Other specific deliverables to be produced during the course of the project will be:

- a report on existing schema currently in use for storing, archiving and documenting qualitative data
- a report on data structures for key CAQDAS packages and large web-based multi-media projects
- reports from liaison with software suppliers

- a detailed and annotated schema (on Sourceforge)
- transformation scripts from the Exchange Format to Qualidata Online interface and to archival copy (Sourceforge)
- draft transformation scripts from Atlas-ti and possibly other softwares to the Exchange Format (Sourceforge)
- Report from workshops and evaluation
- Brief report on plausibility of web-based data conversion service

## 5. Project Outcomes

List the outcomes you envisage, including their impact on the teaching, learning, or research communities, and what change they will stimulate or enable.

The following outcomes will be sought from the project:

- specification of data exchange models and partner consultation;
- pilot conversion tools and testing with partners;
- evaluation of performance and results.

The main outcome is a proposed standard that maintains the integrity of qualitative from a research study – documentation of the study, objects created and content level description (eg text mark-up), linkages between data, and data annotations that others can utilise. This is essential for longer term preservation, data conversion and data portability. Data archivists, data creators, analysts and anyone accessing data now and in the future will benefit from the work.

## 6. Stakeholder Analysis

List key stakeholder groups and individuals that will be interested in your project outcomes, will be affected by them, or whose support/approval is essential, both within your institution and in the community, and assess their importance (low/medium/high).

Stakeholder	Interest / stake	Importance
Researchers across all domains who undertake data collections and analysis using digital formats	More research data that is tied up in proprietary or legacy formats will become available. Data and analysis, rather than simply a single dataset will be able to be long-term preserved. Incorrect conversion of survey data will no longer be necessary. Appreciation of metadata will be improved	Very high
Post graduate students	As above	Very high
Undergraduate students	Low use of primary data in teaching is an issue of concern for some of the Research Councils. More availability of data will be an asset.	Medium
Research funding agencies	Greater value for money from grants awarded through an increase in the archiving and availability of primary data. Increased re-use and further analysis of datasets whose creation they have funded. Greater accountability for public funds through increase in data usage and citation.	High
Library community	Increased recognition by academic and research communities of their libraries' new role in promoting and pointing users to raw data sources for research.	Medium
General public	Typically the general public do not create or analyse research data.	Low

## 7. Risk Analysis

List factors that could pose a risk to the project's success, assess their likelihood and severity, and how you will prevent them from happening (or manage them if they occur). Cover the types of risks listed and any others that apply.

Risk	Probability (1-5)	Severity (1-5)	Score (P x S)	Action to Prevent/Manage Risk	Response to worst case
Staff recruitment	1	5	5	Immediate widescale advertisement of positions subject to award letter. Head-hunting initiated prior to this.	Development timeline reassessed if time permits
Financial mismanagement	1	5	5	Part-time Project Manager and PIs with financial responsibility and project meetings	Institution liable to underwrite losses. Project halted.
Legal problems	1	5	5	Project Manager and PIs well-briefed and sensitive to legal environment.	Advice sought from JISC Legal and ultimately from legal personnel at the University of Essex.
Technical problems	2	5	10	Established track record in technical development. PIs to apply rigorous monitoring.	Technical development strategy reassessed and new staff recruited, if time and resource permits. Development timeline reassessed and new staff recruited, if time and resource permits. New testers recruited.
Staff retention	1	4	4	PIs to monitor	
Tester drop out	1	5	5	PIs to monitor	

## 8. Standards

List the standards the project will use in the table below. Also indicate:

- Any deviations from the standards that JISC recommends.
- Where choices exist in an area, the reasons for the standards selected.
- Where proprietary standards are selected in an area where open ones are available, the reasons for their use and their scope of deployment.

Name of standard or specification	Version	Notes
TEI		
DDI	3	
Dublin Core		
QDIF	1	
METS		
SDMX		
SSS		

## 9. Technical Development

Indicate how the project will follow best practice for technical development, and any specific technologies or development approaches the project will adopt and why.

The Project will adhere to the *JISC Information Environment Architecture Standards Framework*. Standards and protocols employed will be open wherever possible. These are likely to include Java, OpenURL, OAI-PMH, SRU/W, XML and SQL. Metadata will conform to Qualified Dublin Core. UKDA uses the Data Documentation Initiative (DDI) and the Text Encoding Initiative (TEI), which both map

onto Dublin Core but provide more relevant detailed description. METS will most likely be the first choice for overarching schema for WP3.

## 10. Intellectual Property Rights

*Indicate who will own the intellectual property created by the project List any intellectual property owned by third parties that will be incorporated into project outputs, when/how you will obtain permission to use them, and any implications for project outputs after the project ends.*

Software developed will be made available to the UK HE/FE community on an open source basis. Where appropriate, all outputs will be made available via SourceForge for further development.

This project seeks to undertake groundwork preparatory to further infrastructure development. It will conclude with recommendations for a greater promotional effort towards uptake of the data standards and a full-scale data conversion facility/service, and the UKDA will archive the project web site for as long as JISC requires, up to a maximum of five years after completion of the project.

## Project Resources

### 11. Project Partners

*List all project partners (including subcontractors), their roles, and the main contact. Indicate the date a consortium agreement was signed (or will be signed), and send a copy to the programme manager.*

This is a small project and as such is limited to a single institution. A number of organisations in England and Wales have agreed to be user testers.

### 12. Project Management

*Briefly describe the project management framework, including organisation, reporting relationships, decision process, and the role of any local management committee.*

*List all members of the project team, their roles, and contact details. Indicate the proportion of time the project manager will spend on project management.*

*Indicate if the project has training needs and how they will be met.*

Project management for this project will ensure that all the workpackages of the project are managed coherently and that all the project outputs are delivered within the agreed deadlines and budget.

The staff committed to this work are:

1. Project Director, Louise Corti - .2 FTE who will be ultimately responsible for all the elements of the project, and will lead WP3. She will also be in charge of producing workplans, running project meetings, production of bi-monthly progress reports and key reporting to JISC, oversight of project web site (This post is being contributed at no cost to the JISC).
2. Investigator for WP2, Matthew Woollard - .1 FTE, who will lead and oversee the work for WP2 and contribute relevant parts to reports as requested.
3. Technical Officer, Angat Bhat, 1FTE, leading the technical work for WP3
4. Project administrator, Hilary Beedham – .05 FTE who is responsible for contract and financial management
5. Clerical Officer, Claire Flaxton – 0.1 FTE, response for general clerical support

A small Project Committee will be formed to steer and advise the project. The Project Committee will be composed of people with data curation or data handling expertise in the UK who will meet 3-4 times a year (face-to-face, VC or AGN) to consider progress, advise on quality assurance and provide advice based on leading edge ideas in their fields. They will also help to disseminate the work of the project, particularly among senior staff in UK universities and research agencies, to government, and overseas. The Board will consist of 'research community champions' from the domains represented in

our project, together with senior figures from the worlds of digital libraries and data curation. The members are:

- UK Data Archive, Ken Miller, Head of Information Development and e-Social Science;
- UK Data Archive, Matthew Woollard, Head of Digital Preservation and Systems;
- Data Curation Centre (DCC), Chris Rusbridge
- A representative from the statistical community (tba);
- University of Queensland - Andrew Smith.

### 13. Programme Support

Indicate if there are specific areas where you would like support from the programme or programme manager.

No

### 14. Budget

Use the budget template and attach the project budget as Appendix A. Explain any changes from the budget in the agreed project proposal.

## Detailed Project Planning

### 15. Workpackages

Use the workpackages template to plan the detailed project work and attach as Appendix B. Clearly indicate project deliverables and reports (in **bold**), when they are due, phasing of workpackages, and explain any dependencies. You may also attach a Gantt chart, diagram, or flowchart to illustrate phasing.

### 16. Evaluation Plan

Indicate how you will evaluate the quality of the project outputs and the success of the project. List the factors you plan to evaluate, questions the evaluation will answer, methods you will use, and how success will be measured. Expand as appropriate on how you will conduct the evaluation.

Formative evaluation will be the responsibility of the PIs

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Mths 1-12	WP3 Use of metadata/data standards	Which ones to use?	Fitting model data to schema and schema validation	Validation success
Month 5	WP3 Suitability of model for software exchange with suppliers	Probability of buy in to model; import and export requirements	Face to face group meeting	% Take up from software suppliers
Months 3-12	WP3 front end to schema import and export tools	Ease of use and flexibility	Testing on data processing in UIKDA	% positive review
Mths 1-6	WP2 Use of metadata/data standards	Which ones to use?	Fitting model data to schema and schema validation	Validation success
Months 6-9	WP2 front end to schema import and export tools	Ease of use and flexibility	Testing on data processing in UIKDA	% positive review

Month 6	Progress of all approaches, scema and tools	Realistic?	Advisory Committee	% positive review
---------	---	------------	--------------------	-------------------

## 17. Quality Plan

Explain the quality assurance procedures you will put in place to ensure that project deliverables meet quality expectations and acceptance criteria. Complete the table below for each of the major deliverables providing as much detail as possible. Repeat the table as many times as necessary to accommodate all deliverables.

Output					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibilities	Quality tools (if applicable)
Month 12	Documentation	Production of understandable documentation for schema and tools	Can be understood by non technical staff		n/a
Month 11-12	Usability	Testing on processing staff	Can be used by non technical staff		n/a
Month 12	Accessibility	W3C validation, Sourceforge deposit	Validated, formal deposit		Schema validaton

documentation, usability and accessibility

## 18. Dissemination Plan

Explain how the project will share outcomes and learning with stakeholders and the community. List important dissemination activities planned throughout the project, indicating purpose, target audience, timing, and key message.

Timing	Dissemination Activity	Audience	Purpose	Key Message
Month 3	Website	All	Information and publicity	Dext information
Month 5	WP3 Meeting	All CAQDAS Software Suppliers	Present schema and get buy in	Buy in
Month 5	WP3 Conference presentation	CAQDAS users	Present schema and get buy in	Buy in
Month 6	WP3 Conference presentation	IASISST Data archiving community	Knowledge giving	Provide a standard
Month 7	Leaflet	All	Knowledge giving and promotion	DExT information
Month 7	Advisory Committee review of draft schema and tools	AC	Review, OK	DExT strategy and progress
Month 10	WP3 and WP2 Conference presentation	ASC Survey methods and data community	Knowledge giving	Provide a standard
Month 10	Workshop	Potential users	Presenting ideas, promotion, buy in	Here is standard and tools
Month 11	Schema and tools Documentation	All users	Guide to standard and tools	Guide to standard and tools

## 19. Exit and Sustainability Plans

*Explain what will happen to project outputs at the end of the project (including knowledge and learning). Focus on the work needed to ensure they are taken up by the community and any work needed for project closedown, e.g. preservation, maintenance, documentation.*

<b>Project Outputs</b>	<b>Action for Take-up &amp; Embedding</b>	<b>Action for Exit</b>
Schemas	Sourceforge and promotion to data archivists and software suppliers. Longer term use within UKDA	All on SourceForge and UKDA Website. Incorporate into UKDA practise
Tools	Sourceforge and promotion to data archivists and software suppliers. Longer term use within UKDA	All on SourceForge and UKDA Website. Incorporate into UKDA practise

*List any project outputs that may have potential to live on after the project ends, why, how they might be taken forward, and any issues involved in making them sustainable in the long term.*

<b>Project Outputs</b>	<b>Why Sustainable</b>	<b>Scenarios for Taking Forward</b>	<b>Issues to Address</b>
See above			

Appendixes

## **Appendix A. Project Budget**

### **JISC Project Budget Template**

*Select the appropriate template and attach to the proposal, plan, report, etc.  
Fill in the information for the header, e.g. project acronym, version, and date.*

### **Template for Project Proposal and Project Plan**

## Appendix B. Workpackages



WORKPACKAGES	Month	1	2	3	4	5	6	7	8	9	10	11	12
<b>1: Project Management</b> <i>Objective:</i> to ensure that all the workpackages of the project are managed coherently and that all the project outputs are delivered within the agreed deadlines and budget													
<b>2: Develop technical specification of data exchange models – survey data</b>													
<b>3: Develop technical specification of data exchange models – qualitative data</b>													
<b>4: Testing and Evaluation</b> <i>Objective:</i> to ensure that formative evaluation takes place throughout the project and is reported to JISC, and a 2 month final summative evaluation is performed to consider recommendations of the project.													

Project start date: 11 December 2006  
 Project completion date: 10 December 2007  
 Duration: 12 months

Workpackage and activity	Earliest start date	Latest completion date	Outputs (clearly indicate deliverables & reports in bold)	Milestone	Responsibility
<b>YEAR 1</b>					
<b>WORKPACKAGE 1:</b>					
<b><u>Objective:</u></b>					
1. Ensure that all the workpackages of the project are managed coherently	Month 1	Month 12	Workplan Meeting notes Progress reports		PD PD PD/PIs
2. Ensure that all the project outputs are delivered within the agreed deadlines and budget	Month 1	Month 12	Progress reports 6 months		PD
3. Contract management	Month 1	Month 12	Contract		PA
4. Financial management	Month 1	Month 12	Returns		PA
5. Website	Month 3	Month 12	W3C compliant website		PD
6.					
<b>WORKPACKAGE 2:</b>					
<b><u>Objective:</u></b>					
1. Technical specification of data exchange model for survey data	Month 1	Month 12	Report		PI2/TO2
2. Production of a metadata model	Month 1	Month 7	Schema, Source Forge		TO2
3. Development of tools sets for conversion	Month 3	Month 11	Tools, Documentation, Sourceforge		TO2
4. Development of tools sets for archiving	Month 10	Month 12	Tools, Documentation, Sourceforge		TO2
<b>WORKPACKAGE 3:</b>					
<b><u>Objective:</u></b>					
1. Technical specification of data exchange model for survey data	Month 6	Month 12	Report		PI3/TO3
2. Production of a metadata model	Month 6	Month 11	Schema, Source Forge		TO3
3. Development of tools sets for conversion	Month 8	Month 12	Tools, Documentation, Sourceforge		TO3

4. Development of tools sets for archiving	Month 10	Month 12	Tools, Documentation, Sourceforge		TO3
<b>WORKPACKAGE 4:</b>					
<b><u>Objective:</u></b>					
1. ensure that formative evaluation takes place throughout the project and is reported to JISC	Month 1	Month 12	Progress reports Feedback		
2. Summative evaluation – meeting	Month 5		WP3 Workshop material and Feedback report		PD/PI2 and TO2
3. Summative evaluation – committee and partners	Month 9	Month 12	Usability test reports		PD/PIs and TOs
4. Summative evaluation - workshop	Month 10	Month 12	Usability test report		PD/PIs and TOs
5. Summative evaluation – Sourceforge activity	Month 12	Month 12	Report		PD/PIs and TOs

Members of Project Team:

PD project Director  
 PI2 Principal Investigator WP2  
 PI3 Principal Investigator WP3  
 PA Project Administrator  
 TO2 Technical Officer WP2  
 TO3 Technical Officer WP3

---

<sup>i</sup> Draft qualitative data DTD [www.esds.ac.uk/qualidata/online/about/draft.asp](http://www.esds.ac.uk/qualidata/online/about/draft.asp)