

# JISC DEVELOPMENT PROGRAMMES

## Project Document Cover Sheet

### Project

<b>Project Acronym</b>	-	<b>Project ID</b>	
<b>Project Title</b>	19 <sup>th</sup> Century Pamphlets Online		
<b>Start Date</b>	1 March 2007	<b>End Date</b>	28 February 2009
<b>Lead Institution</b>	University of Southampton		
<b>Project Director</b>	Dr Mark Brown, University Librarian		
<b>Project Manager &amp; contact details</b>	Grant Young ILRT University of Bristol 8-10 Berkeley Square Bristol B8 1HH Phone: +44 117 9287109 Grant.Young@bristol.ac.uk		
<b>Partner Institutions</b>	BOPCRIS/University of Southampton; University of Bristol; CURL; Durham University; JSTOR; University of Liverpool, LSE; University of Manchester; MIMAS; University of Newcastle; UCL		
<b>Project Web URL</b>	<a href="http://www.curl.ac.uk/projects/NineteenOnline.htm">http://www.curl.ac.uk/projects/NineteenOnline.htm</a>		
<b>Programme Name (and number)</b>	Digitisation Programme		
<b>Programme Manager</b>	Alastair Dunning		

### Document

<b>Document Title</b>	Project Plan		
<b>Reporting Period</b>			
<b>Author(s) &amp; project role</b>	Grant Young, Project Manager; Julian Ball, Technical Project Manager; Mark Brown, Project Director; Richard Wake, Budget Manager		
<b>Date</b>	31 May 2007	<b>Filename</b>	PamphletsProjectPlan01
<b>URL</b>			
<b>Access</b>	<input checked="" type="checkbox"/> Project and JISC internal	<input type="checkbox"/> General dissemination	

### Document History

<b>Version</b>	<b>Date</b>	<b>Comments</b>
01.0	31 May 2007	
_FOI	28 Nov 2007	FOI copy with details of budget, licensing agreements and email addresses excluded

# 19<sup>th</sup> Century Pamphlets Online – Project Plan

## *Overview of Project*

### 1. Background

This is a two-year project led by the University of Southampton on behalf of CURL (Consortium of Research Libraries in the British Isles). It represents the first phase of what is hoped will be a much larger project. That larger project has the vision of providing researchers, teachers and learners with online access to the most significant collections of 19<sup>th</sup> century pamphlets held within UK research libraries.

This current project will draw on the collections of seven libraries, focusing particularly on pamphlets that address the political, economic and social issues that fuelled the great debates and controversies of the 19<sup>th</sup> century, both in Parliament and across the country.

#### **1.1 Opening up a valuable but underused resource**

Pamphlets played an important role within 19<sup>th</sup> century political discourse, representing diverse contemporary perspectives, often polemical or propagandist in nature. They are a valuable primary source that can complement other sources, such as newspapers, periodicals or parliamentary papers. However because of their ephemeral nature, they are often very scarce and difficult to access and are consequently underused within research and teaching.

From 1999-2002 a large retrospective cataloguing project, sponsored by the Research Support Libraries Programme (RSLP) and CURL, catalogued nearly 180,000 19<sup>th</sup> century pamphlets from 21 research libraries<sup>1</sup>. That project has greatly assisted researchers in finding pamphlets via local library catalogues and the combined academic and national library catalogue Copac<sup>2</sup>.

However, having discovered the existence of a pamphlet, a researcher will then often face the barrier of having to travel to a distant library to view the item, since 19<sup>th</sup> century pamphlets are usually held within special collections and seldom loaned out. Although a published item, their typically short print-runs and ephemeral nature mean that 19<sup>th</sup> century pamphlets can now be very scarce. A scoping study undertaken in preparation for the final project bid checked a random sample of approximately 100 pamphlets from each of the seven contributing libraries against the Copac catalogue. It found that between 23% and 44% of the pamphlets were only recorded as being held by that library<sup>3</sup>. Peter Scott, a researcher at the University of Reading, has described his experience:

---

<sup>1</sup> See at <http://www2.is.bham.ac.uk/rslp/pamphlets/pamphlets.htm>

<sup>2</sup> See at <http://copac.ac.uk/>

<sup>3</sup> Young, G., *19<sup>th</sup> Century Pamphlets Online: digitisation scoping study*, (JISC, 2006), p.13.

Available online at:

[http://www.jisc.ac.uk/publications/publications/pub\\_digi\\_scopingstudy.aspx](http://www.jisc.ac.uk/publications/publications/pub_digi_scopingstudy.aspx) [Accessed 3 May 2007]

Pamphlets are an important but under-utilised historical resource. I frequently use Copac to track them down but am then faced with time-consuming and expensive journeys to look at rather short documents.<sup>4</sup>

It is likely that only the most determined would go to such trouble – or those who already have access to significant pamphlet collections. Where researchers, teachers or learners have such access, pamphlets are often highly valued. To quote Claire Bronze, a medical student from UCL:

...my dissertation was focused on eighteenth and nineteenth century medical history, specifically looking at anatomy teaching, human dissection and grave-robbing in this period. I referred to the Hume Tracts collection [the UCL collection being contributed to this project] for much of the research in my project and they were very helpful in allowing me to achieve a detailed picture of the politics, scientific advances and public perception of this field of medicine. Having primary evidence of this sort helped me to construct a convincing argument as to the situation surrounding anatomy teaching at this time.<sup>5</sup>

Durham University, another contributing library with a significant collection, have made much use of their pamphlets in outreach to schools: using them with 11-14 year olds to investigate issues of bias, truth and reliability in historical evidence, and with six formers to chart the progress of political reform as part of the history curriculum.

By linking digitised pamphlets to the existing catalogue records created under the RSLP/CURL project, this project takes a further and vital step: ensuring that researchers, teachers and learners do not just discover the existence of such items, but can access them without difficulty.

## **1.2 Ensuring wide resource discovery, access, and use**

The partnerships and technologies the project puts in place will ensure that resource discovery and access is even broader than local library catalogues or the combined Copac catalogue, as important as these are.

In addition to cataloguing records, the previous RSLP/CURL project created an online *Guide to 19<sup>th</sup> Century Pamphlets*<sup>6</sup>, which provided collection-level descriptions of the collections that were catalogued. This current project will re-develop the *Pamphlets Guide* to enable collection-based and subject-based access for the digitised pamphlets via the Copac database.

The project places the digital collection itself with the scholarly archive JSTOR<sup>7</sup>, and so is also able to take advantage of JSTOR's search and retrieval interface, marketing activities, and linking arrangements with other organisations. JSTOR has arrangements with Google, the History Cooperative<sup>8</sup>, and RePEc (Research Papers in Economics)<sup>9</sup> among others. The pamphlets will appear within Google Scholar<sup>10</sup> and be fully indexed by Google's spider, enabling them to be found via a standard

---

<sup>4</sup> Quoted in this current project's final bid.

<sup>5</sup> Also quoted in final bid.

<sup>6</sup> See at <http://www.curl.ac.uk/rsllpguide/guidehp.htm>

<sup>7</sup> See at <http://www.jstor.org/>

<sup>8</sup> See at <http://www.historycooperative.org/>

<sup>9</sup> See at <http://repec.org/>

<sup>10</sup> See at <http://scholar.google.com/>

Google web search<sup>11</sup>. Through JSTOR's participation in CrossRef<sup>12</sup>, references to the pamphlets in many scholarly articles will carry links through to the digitised pamphlets.

JSTOR will also seek to create rich links between items in the pamphlet collection and its other collections. Whereas the CrossRef service enables an article referencing a pamphlet to link directly to that pamphlet, the interlinking proposed by JSTOR is likely to work both ways – enabling, for example, pamphlet users to link to articles about the authors of the pamphlet or the issues it addresses. This richer linking would only be available to those subscribing to the relevant collections, but for the many UK institutions that do it is likely to add considerable value to the collection.

While the pamphlet collection as a whole will be delivered by JSTOR, the libraries contributing to the project will also be entitled to take copies of the digitised pamphlets for their own local delivery and use. Other UK users will be able to download individual pamphlets from JSTOR and incorporate them within their teaching and learning resources for local use or for contribution to shared learning repositories such as JORUM<sup>13</sup>. This local use and reuse of the collection will provide another means for people to encounter the pamphlets and the project will seek to encourage such use by providing an exemplar learning resource for JORUM.

### **1.3 Building on previous digitisation experience**

The digitisation itself will be undertaken by BOPCRIS<sup>14</sup>, a specialist digitisation laboratory based within the University of Southampton's library. BOPCRIS have considerable experience in digitising historic textual material, most recently in the digitisation of a million pages of 18<sup>th</sup> century parliamentary papers for Phase 1 of the JISC's Digitisation Programme.<sup>15</sup> This current project takes advantage of the hardware, software, and expertise accumulated in BOPCRIS's previous work.

### **1.4 Ensuring sustainability and digital preservation**

Two key challenges for large digitisation projects are to ensure that (1) the access they provide to resources is financially sustainable and (2) the data itself is preserved as formats, platforms and other technologies change. This current project addresses both issues through its business model, in particular through the partnership with JSTOR.

JSTOR will preserve a copy of the archival digital dataset created by BOPCRIS from the library collections. It will undertake all the activities required to preserve this dataset, including backing-up, data checking, and migration to other formats. Contributing libraries will be able to request copies covering their own collections, while CURL, the JISC or HEFCE will be able to request a copy of the entire dataset.

JSTOR will also derive a delivery dataset, which it will mount as a JSTOR collection and make freely available to UK secondary schools, FE, HE and some other

---

<sup>11</sup> See at <http://www.google.co.uk/>

<sup>12</sup> See at <http://www.crossref.org/>

<sup>13</sup> See at <http://www.jorum.ac.uk/>

<sup>14</sup> See at <http://www.southampton.ac.uk/library/about/projects/bopcris.html>

<sup>15</sup> See at <http://www.bopcris.ac.uk/18c/>

institutions. The sub-licence for UK users will enable items from the collection to be repurposed for research, teaching and learning applications and included within learning or research repositories.

The costs of archiving and delivering the collection for UK users will be entirely borne by JSTOR and will be funded through income it is able to generate in other markets. Underpinning the business model is an agreement between CURL (representing UK partners) and JSTOR, which will have a 25-year term (with appropriate review points and the option of continuing beyond the 25 years).

As this background section has indicated, the project builds on previous work and takes advantage of existing infrastructure. The approaches it adopts are intended to maximise the return on investment by avoiding some of the more costly elements involved in developing a digital collection. In particular the project will:

- Avoid expensive metadata creation by utilising the existing high-quality catalogue records created under the recent RSLP/CURL project.
- Avoid considerable infrastructure, equipment, and training costs by centralising the scanning within the well-equipped BOPCRIS laboratory.
- Avoid the costs and risks associated in developing a standalone delivery system by making use of JSTOR's existing delivery platform.
- Exploit existing, multiple resource discovery services, including Copac, JSTOR, Google, and the CrossRef initiative.
- Take advantage of JSTOR's marketing and publicity activities and its interlinking with other collections.
- Ensure the preservation and sustainability of the collection through its agreement with JSTOR.

## 2. Aims and Objectives

The **overall aim** of the project is:

**To provide researchers, teachers and learners with online access to significant collections of 19<sup>th</sup> century pamphlets held within UK research libraries.**

In order to achieve this aim, the project has **five key objectives**:

- 1. To digitise a wide selection of 19<sup>th</sup> century pamphlets focusing on political, social and economic issues.**  
The project intends to deliver 1 million pages of 19<sup>th</sup> century pamphlet literature (approximately 23,000 pamphlets) along with associated OCR text and metadata. The pamphlets will be drawn from the collections of seven research libraries and provide both breadth and depth of coverage, whilst focusing particularly on political, social and economic themes.
- 2. To establish an efficient consortial scanning operation.**  
The centralised scanning at BOPCRIS will provide savings but require some careful coordination and management, particularly in avoiding duplication, ensuring a steady flow of material, and assuring high quality outputs. The project will seek to put in place an efficient and quality-assured workflow that will serve both this current project and future work.
- 3. To provide sustainable preservation and delivery.**  
Through its partnership with JSTOR, the project seeks to ensure that the collection is preserved and delivered free to UK secondary schools, FE and HE for at least 25 years.
- 4. To enable sophisticated, distributed resource discovery and access.**  
Through its partnerships with JSTOR and MIMAS, the project seeks to provide multiple means of resource discovery and access, utilising existing infrastructures and both scholarly and popular search services.
- 5. To provide models for further phases/projects.**  
The project will seek to capture and disseminate the key lessons and models from this project that can be applied to further phases or to other, similar projects.

## 3. Overall Approach

### 3.1 Strategy / Methodology

This section outlines some of the main tasks being undertaken by the project, indicating strategies and methods as appropriate (some are discussed in more detail in later sections of this plan).

#### 3.1.1 Selection and preparation of the pamphlets (Libraries)

Because it is not practical to individually select 23,000 pamphlets for inclusion in such a project, the decision was made during the bid preparation to identify several collections that could be scanned in their *near entirety* (i.e. excluding non-19<sup>th</sup> century, in-copyright, or fragile/incomplete pamphlets). Apart from being pragmatic, this approach provides an extra dimension to the digitised collection: enabling the pamphlets to be understood within the context of particular historic collections. Such an approach will be facilitated through the collection-level *Guide to 19<sup>th</sup> Century Pamphlets*<sup>16</sup> and by provenance information included within the pamphlet's metadata.

Many collections were put forward for consideration by CURL member libraries. Complete collections from Durham, Liverpool, Manchester, Newcastle and UCL were chosen, to provide a wide and balanced range of content. In addition, the bid team decided to make a *selection* of 19<sup>th</sup> century pamphlets from two of the UK's larger pamphlet collections: Bristol and LSE. Both collections have a strong political emphasis. The selection from these libraries will aim to highlight the strengths of their collections and also seek to fill in any obvious gaps left by the other collections.

The table below provides an overview of the collections and the approximate number of pamphlets they will contribute to the project. Fuller details for these collections can be found in the project's *Scoping Study*<sup>17</sup> (pp.15-20) and the RSLP/CURL *Pamphlets Guide*.

---

<sup>16</sup> See at <http://www.curl.ac.uk/rsplguide/guidehp.htm>

<sup>17</sup> All further references to the *Scoping Study* refer to: Young, G., *19<sup>th</sup> Century Pamphlets Online: digitisation scoping study*, (JISC, 2006), p.13. Available online at: [http://www.jisc.ac.uk/publications/publications/pub\\_digi\\_scopingstudy.aspx](http://www.jisc.ac.uk/publications/publications/pub_digi_scopingstudy.aspx) [Accessed 3 May 2007]

Figure 3.1:A. Pamphlet collections

Contributing Libraries	Collections	Pamphlets (estimate*)
Durham	<b>Earls Grey</b> <i>Collection belonging to family of politicians and colonial administrators</i>	1,160
Liverpool	<b>Earls of Derby (Knowsley)</b> <i>Collection belonging to family of politicians and colonial administrators</i>	1,209
UCL	<b>Joseph Hume (1777-1855)</b> <i>Personal collection of an MP; predominately 1<sup>st</sup> half of 19<sup>th</sup> century</i>	3,528
Newcastle	<b>Joseph Cowen (1829-1900)</b> <i>Personal collection of an MP; predominately 2<sup>nd</sup> half of 19<sup>th</sup> century</i>	1,579
Manchester	<b>Foreign Office &amp; Colonial Office</b> <i>Government collections focused on international relations and the empire</i>	3,149
Bristol	<b>Selection from pamphlet holdings</b> <i>Particularly from the National Liberal Club collection, which includes personal and party collections</i>	6,000
LSE	<b>Selection from pamphlet holdings</b> <i>LSE is strong in party and pressure-group collections</i>	6,000
Total		23,125
*These figures are from the scoping study and attempt to take into account duplication. The actual number will vary depending on the order in which the collections are scanned.		

For the five complete collections the project has devised a *de-selection* strategy. This is outlined in the *Scoping Study* (pp.21-22). De-selection is made for one of four reasons:

1. **Pamphlet is in copyright**
2. **Pamphlet is published outside the bounds of the 19<sup>th</sup> century** (we have permitted a small proportion of late 18<sup>th</sup> century material, but will not allow pamphlets published after 1900)
3. **Pamphlet has already been digitised** (or sent for digitisation) from another collection
4. **Pamphlet is incomplete or too fragile** to digitise

The project will create a database (referred to elsewhere as Library Partners' Database) and associated workflow in order to help libraries manage the de-selection process. The database will be pre-loaded with records from Copac for each

collection, enabling libraries to easily identify duplicates and record information about the copyright status and physical condition of their own items. The *Scoping Study* presented a workflow based on the use of this proposed database (pp.24-26). A revised version is presented in Appendix C below.

For Bristol and the LSE a *selection* of pamphlets will be made. This will be made by library staff familiar with the collections, but will be influenced by the following criteria:

1. **Their relevance to themes of the great 19<sup>th</sup> century debates** (e.g. universal suffrage, relationship of church and state, colonial policy) as identified by the collection curators and the academics and teachers participating in the Project Steering Group.
2. **Their usefulness in addressing gaps** in the digital collection (e.g. themes not well covered, formats not represented, particular authors who should be included). Again, the project will take advice from the Project Steering Group.
3. **Feedback and demand from collection users.** Consignments from Bristol and the LSE will be scheduled near the end of the project, by which stage the project aims to have material online and the possibility to track usage and survey some users.
4. **Replacements** for copies held in the smaller collections that are incomplete or too fragile to digitise.

Again, the Library Partners' Database will be used by Bristol and the LSE to assist in identifying potential pamphlets within their own collections, and to identify duplicates and damaged items in other collections. As with the other libraries, Bristol and the LSE will use it to identify items they are sending, record their copyright status, and make any relevant notes about the physical condition of the pamphlets.

Some of the collections have all their pamphlets bound in volumes (Liverpool, LSE, Newcastle, UCL); some have a mixture of separate pamphlets and volumes (Bristol, Manchester); and one has separate pamphlets only (Durham). This means that in some cases libraries will be sending volumes that include pamphlets not intended for scanning. Depending on the library's preference, items scheduled for scanning are denoted by either slips within the volumes (containing the Copac identifier used to link it to the database and metadata record) or a printed list for each volume indicating which items are to be scanned (this can be generated from the Library Partners' Database).

Despite the use of the database, pamphlet preparation will be time consuming. The project aims to start the preparation sufficiently in advance of the scheduled delivery dates to ensure that the pamphlets are ready for transport to Southampton. Appendix B includes the current schedule for library preparation and collection deliveries (see Workpackage 4). It is possible that this may have to be revised as information about the preparation time becomes available. In recognition of the work they are undertaking in for this project, libraries have been allocated £1.50 per pamphlet scanned. It is anticipated that they will be committing at least that amount again in kind (including preparation of pamphlets, checking them upon their return, and other costs associated with the project such as meeting attendance).

### **3.1.2 Creation of archival digital dataset (BOPCRIS)**

The digitisation is centralised at BOPCRIS within the University of Southampton Library. An MOU between BOPCRIS and each partner library covers transportation, handling, reporting and insurance of the materials while in transit and at the University of Southampton.

BOPCRIS receives consignments of physical pamphlets from the libraries, checks and then formally 'receives' them using the Library Partners' Database. Damage to any of the volumes or individual pamphlets that has been noted by the partner libraries is acknowledged within the database by BOPCRIS. Any additional damage identified by BOPCRIS is also noted and conveyed to the libraries. The transportation, checking, logging of pamphlets, and their return to the partner libraries will be managed by the Inventory Project Officer.

A scanning methodology will be agreed between BOPCRIS and JSTOR for the pamphlet collection. Scanning staff and the QA Project Officer will produce images of each pamphlet page and from those images produce a text equivalent using OCR. Metadata will be generated in XML format for each page and this will be in part derived from the Copac record retrieved from the Library Partners' Database. Section 8 of this plan describes the standards to be used for each of these digital files.

All stages of the workflow within BOPCRIS and between BOPCRIS and JSTOR will be managed by a custom built database with a web interface (see section 9 below).

The digitisation workflow at BOPCRIS follows this order:

1. The Project Officers determine the scanning requirements for each pamphlet or volume.
2. Scanning staff digitise each pamphlet page, undertaking a number of quality control checks through the process. An overall page count check is made at the completion of each pamphlet.
3. Within an XML editor, Project Officers import the XML, carry out a final check that all pamphlet pages have been scanned, and ascribe the logical pagination. Foreign texts and old English typefaces are tagged in the xml.
4. The completed XML is then imported into the OCR workflow to generate the OCR txt for each pamphlet page. This stage also includes deskewing or despeckling, if required. The presence of Foreign languages or old English typefaces triggers the usage of specialist OCR software that can better transcribe these.
5. An XML record is generated, incorporating the MODS record and MIX attributes within a METS XML wrapper.
6. The tiff images, OCR text files and XML metadata are supplied to JSTOR for checking, archiving, and use in deriving a delivery dataset.
7. A further metadata file is supplied to JSTOR for archiving that will include PREMIS metadata. The PREMIS attributes and protocol will be determined during the initial phase of the project.

### **3.1.3 Preservation and delivery of the collection (JSTOR)**

The datasets are transferred by BOPCRIS to JSTOR's FTP site in regular batches. JSTOR will pass the data through its own quality assurance process, checking an average of 10% of images and OCR files for accuracy. Where there are any problems, JSTOR will liaise with the Technical Project Manager at BOPCRIS to try to address them. Because some rescanning may be necessary, collections will be held at Southampton until they have been signed off by JSTOR.

JSTOR will store a copy of the archival dataset within their dark archive and preserve it for future use or provision of copies back to contributing libraries, CURL, or JISC/HEFCE. Delivery images, texts and metadata will be derived from the archival dataset for use within JSTOR's own proprietary delivery system. Section 8 describes the standards JSTOR uses in delivering the collection. Content will be released in batches, enabling the project to begin delivering content within, or very soon after, its first year and providing an opportunity for evaluation and to inform the later phase of the selection from Bristol and the LSE.

### **3.1.4 Resource discovery (JSTOR, MIMAS, Project Team, and Libraries)**

As monographic/pamphlet literature is a new form of content for JSTOR, it will work closely with the Project Team to determine the best way to present it to users. The project expects the collection to have a unitary identity for UK users, but it may also be presented as sub-collections or associated with other JSTOR content. We envisage multiple ways into the collection including both browse- and search-based access.

Once the digital pamphlet is ready it will be assigned permanent URLs and Digital Object identifiers (DOIs), which will enable the pamphlets to be linked to via CrossRef and similar initiatives.

JSTOR will also pass the unique identifiers to MIMAS, to enable them to be incorporated within the Copac records. MIMAS will additionally make these records available to CURL libraries holding the pamphlets (whether they are project partners or not), enabling libraries to link to digital versions of pamphlets held within their own collections. In order to facilitate this, MIMAS will develop software to identify likely duplicates across the Copac database.

The Project Team, especially the Project Manager and Learning Resource Developer will work with the subject experts on the Project Steering Group and with MIMAS to incorporate effective collection-based and subject-based searching within the re-developed *19<sup>th</sup> Century Pamphlets Guide*.

## **3.2 Issues to be Addressed**

This section discusses several issues identified in the ITT for the programme or that have arisen so far in the planning and initiation of the project.

### **3.2.1 Copyright and licensing issues**

Some late 19<sup>th</sup> century pamphlets within the collections chosen for this project are likely to be in copyright. The project will exclude items known to be in copyright unless library partners undertake copyright clearance (none have indicated that they

intend to do so). Items of uncertain copyright status will be included on a 'notice and take-down' basis with the University of Southampton assuming the risk for these items (to be confirmed). Sampling undertaken in April 2007 found less than 1% of the pamphlets to be in copyright, so the strategy adopted should not adversely affect the quality or scope of the collection. The project's copyright strategy is described in further detail in section 10 below.

A bigger challenge for this project is the licensing between partners. The project has identified 10 agreements that need to be put in place. At the heart of these is the agreement with JSTOR, which underpins the business model and will inform most of the other agreements. Rather than make each partner a party to this agreement, the project's consortium agreement will nominate CURL as the representative of UK partners in the agreement with JSTOR. The terms of the CURL-JSTOR agreement need to be reflected in the other agreements, so this is likely to delay the consortium agreement (as noted in section 11 below). It will also require some additional resources and support (as noted in section 13). Greater detail about the various licenses required is provided in section 10 (below).

### **3.2.2 Accessibility and Usability**

JSTOR's delivery of the collection will comply with the JISC guidelines and IE standards for accessibility and usability. JSTOR currently meets and exceeds W3C WAI Priority 1 and Section 508 of the Rehabilitation Act (a US requirement). JSTOR have paid much attention to site navigation (including keyboard access keys) and users are encouraged to provide feedback.

JSTOR delivers images of each page rather than text (which is primarily used for retrieval). However, the high-quality PDF versions of the pamphlets will enable enlargement and incorporate OCR text for use by screen reader technologies. Larger TIFF images of the pages can also be downloaded for enlargement or for use with a user's own OCR and screen-reading software.

JSTOR's accessibility practices are described in further detail here:

<http://www.jstor.org/about/accessibility.html>

### **3.2.3 Interoperability**

The standards employed by the project (especially the XML-based metadata and DOI identifiers) will help to facilitate interoperability.

By including links within the Copac database, it will be possible for others to create targeted searches by a variety of means, including Z39.50, OpenURL, SRW/SRU<sup>18</sup>. As described above, JSTOR's partnerships and participation within the CrossRef initiative will also help facilitate interoperability.

The standards used in the archival dataset will additionally offer the potential for a partner to expose records for OAI-PMH harvesting. This is not part of the current plan, but the Project Team are exploring this possibility with JSTOR. As the project proceeds, the team will also seek additional opportunities for sharing the collection, such as the NINES initiative.<sup>19</sup>

---

<sup>18</sup> See <http://copac.ac.uk/interfaces/> for more details about these options.

<sup>19</sup> See <http://www.nines.org/>

### 3.2.4 Evaluation

Evaluation of specific deliverables is described in the evaluation plan in section 16 below. Overall evaluation of the project and its outputs will be provided by an externally commissioned report (see Appendix B, Workpackage 10), which will be completed within the last few months of the project.

### 3.2.5 Preservation

The JISC have asked the project to find a UK repository to preserve the archival dataset in addition to JSTOR's dark archive. This requirement was not identified for the revised bid or its budget so may require some further support and funding (see 13 below).

An initial meeting was held with the Arts and Humanities Data Service (AHDS) Executive on 24 May to determine whether the AHDS might be able to provide such services. The AHDS is willing, but require further information about the dataset in order to determine requirements and potential costs. The JISC will be kept informed of progress and involved in discussions as necessary.

## 3.3 Scope

Section 3.1.1 has described the project's methodology for selection and de-selection. This current section provides further information about the scope of the collection and related deliverables.

- **Period.** The previous retrospective cataloguing project defined its 19<sup>th</sup> century as 1801-1914, although included some pamphlets before this period. This project will take the 19<sup>th</sup> century as 1800-1900 inclusive. It does not extend beyond 1900, because of the greater risk of copyright infringement, but it may include some late eighteenth century pamphlets where their exclusion would create significant gaps in collections (e.g. in a sequence of pamphlets).
- **Place and language of publication.** The project is not limited to British publications, but will include items published in other countries and in languages other than English. The Foreign Office and Colonial Office collections from Manchester include significant proportions of foreign publications, as do many of the other collections. Where a selection is being made (i.e. from Bristol and LSE), the project will concentrate on British publications in English.
- **Subject.** As previously stated, the project is focusing on pamphlets related to political, economic, and social issues. However, the collections of families and individuals will include material that addresses other interests (e.g. religious faith or science and technology). These will not be excluded. Indeed, it is hoped that the collection as a whole will represent something of the full range of pamphlets created, circulated and collected in the 19<sup>th</sup> century.
- **Image reproduction.** The emphasis of this project is on providing access to the intellectual content of the pamphlet rather than reproducing an exact facsimile of each physical item. For this reason it seeks to exclude duplicates found within collections and scans bitonally, unless the presence of grey or

colour warrants greater a bit-depth (tonal representation). However, where significant annotations have been made, these will be captured rather than excluded.

- **OCR text.** Although the goal is not facsimile of individual items, the layout of the text on the page is considered part of its intellectual content and is rendered faithfully. In order to achieve this, the user is presented with images of the pages with the OCR text absent (in GIF, JPEG or TIFF presentations) or in the background (in the high-quality PDF version). The primary purpose of the OCR text is to aid retrieval, although it will be possible to copy or access the text via the PDF for repurposing and accessibility (see 3.2.2 above).
- **Metadata.** Descriptive metadata is limited to that available within the Copac records, although the recent cataloguing project means this is of a very high quality. Structural, technical and preservation-related metadata will be created by the project, by automated means where possible. This will be at least minimally compliant with the chosen standards.

### 3.4 Critical Success Factors

The project's success is particularly dependant on upon the following factors:

1. Timely conclusion of the licence agreements required to enable digitisation and delivery (see section 10 below)
2. Steady flow of pamphlets to BOPCRIS from the contributing libraries (see section 3.1.1 above)
3. Timely delivery of high-quality datasets (images, metadata and OCR text) from BOPCRIS to JSTOR (see 3.1.2-3.1.3 above)
4. Timely release of the pamphlet collection in batches by JSTOR (see 3.1.2 above)
5. Effective incorporation of URL/DOI into Copac records and detection of duplicates (see 3.1.4 above)
6. The collection becoming well known and used by intended audiences

## 4. Project Outputs

### 4.1 Deliverables

The project intends to produce the following tangible deliverables.

1. **Archival dataset** containing at least 1 million page images, associated OCR-text and XML metadata (approximately 23,000 pamphlets<sup>20</sup>), preserved by JSTOR and available to contributing libraries, CURL, JISC or HEFCE.
2. **Delivery dataset** containing images, text and metadata suitable for delivery by JSTOR, maintained by JSTOR and with provision for repurposing by UK users in accordance with a JISC Collections sub-licence.
3. **Online pamphlet collection** with appropriate browse and search access and contextual information, developed and maintained by JSTOR.
4. **Enhanced catalogue records** with links to the digital pamphlets, included within the Copac database and available for libraries participating in Copac to download for their local catalogues.
5. **Redeveloped online Pamphlets Guide** (at <http://www.curl.ac.uk/rslpguide/guidehp.htm>), providing learning content and collection- and subject-based access to the JSTOR collection via the Copac database.
6. **Learning resource** to promote the pamphlets and provide an exemplar for their use, deposited within the JORUM learning repository.
7. **A digitisation resource (“toolkit”)** containing documentation of use to later phases or similar projects, mounted on the project website or made available through a suitable JISC service such as TASI.
8. **Workflow-related software** to support the pamphlet preparations (Library Partners’ Database) and the workflow at BOPCRIS. These are described in section 9 below.
9. **A day conference or dissemination event in the summer of 2008.** This may focus more on the use of the pamphlets in research and teaching *or* on the creation of the collection *or* both – the Project Steering Group will be asked to help the Project Team shape this event.
10. **Contributions to other conferences/workshops/events**, both subject-orientated (political, social or economic history) and digitisation-related.
11. **Contributions to publications**, both subject-oriented (political, social or economic history) and digitisation-related.

---

<sup>20</sup> The project’s target is 1 million pages. The *Scoping Study* estimated this as 23,000 pamphlets based on sampling (p.15), but the actual number achieved by the project may prove higher or lower.

12. **Project documentation**, including but not limited to:

- a. Project Plan (current document)
- b. Licence Agreements (see section 10)
- c. Memoranda of Understanding (see section 10)
- d. Reports to the JISC (progress reports and the final and completion reports)
- e. Internal discussion documents and reports (ad hoc)
- f. Technical documentation (including specifications, guidelines, workflow diagrams)
- g. Minutes of Project Management Group and Project Steering Group meetings

#### **4.2 Knowledge, experience, and other outputs**

In addition to the tangible deliverables listed above, the project expects to gain some further knowledge and experience.

1. **The project hopes to contribute to or facilitate further research into the 19<sup>th</sup> century pamphlet as an historical literary form.** A large corpus of digitised pamphlets will enable, for example, texts to be more easily compared, and the “threads” of various debates to be read through sequences of pamphlets. Similarly, the availability of this collection will enable relationships to be explored with other forms of literature, such as newspapers, periodicals, books and manuscripts. This output might form part of the proposed conference or take the form of presentations and articles within other events and publications.
2. **The project hopes to discover and promote the potential of 19<sup>th</sup> century pamphlets for use within teaching and learning contexts** at secondary and tertiary levels. This output is likely to influence the development of the learning content proposed for the *Pamphlets Guide* and for JORUM.
3. **The project will check the findings and approaches of the *Scoping Study*.** The project’s *Scoping Study* recommended that the project should “evaluate the findings and approaches chosen in this study in the light of the practical reality of the project, and disseminate its findings” (p.45). Examples of these are the sampling methods used in the scoping study to estimate page averages and duplication, or the factors suggested as impacting on scanning time and quality (e.g. in-margin stitching). This knowledge and experience would be incorporated within the digitisation toolkit and reported on in contributions to conferences and publications. It will also inform future work undertaken by BOPCRIS.
4. **The project expects to learn much in implementing the metadata standards it has chosen.** METS and MODS are established standards, although their implementation is not yet straightforward (particularly in the case of METS). MIX and PREMIS are newer and the project will be an early adopter. Again, the lessons learned from this work are likely to be incorporated within the digitisation toolkit and reported on in contributions to conferences and publications. They will also inform future work undertaken by BOPCRIS.

## 5. Project Outcomes

The project envisages the following outcomes:

1. Greater ease of discovery, access and use of pamphlets.
2. More awareness and use of pamphlets within 19<sup>th</sup> century research and teaching.
3. Opportunity for further scholarly research into the pamphlet form with the availability of a large accessible digital archive.
4. Bringing together of disparate items, making it easier to study an individual's or organisation's body of work and to follow individual debates through a sequence of pamphlets.
5. Reduced wear and tear on some fragile original copies.
6. Greater use of original pamphlets not included in this phase of the project.
7. Testing of an alternative business model for resource preservation and sustainability.
8. Testing of an alternative, distributed approach to resource discovery.
9. Testing of newer metadata standards (especially MIX and PREMIS).

## 6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
JISC	Major funder	High
CURL	Sponsor and funder	High
University of Southampton / BOPCRIS	Grant holder	High
University of Bristol	Project Partner	Medium/High
Durham University	Project Partner	Medium/High
University of Liverpool	Project Partner	Medium/High
LSE	Project Partner	Medium/High
University of Manchester	Project Partner	Medium/High
University of Newcastle	Project Partner	Medium/High
UCL	Project Partner	Medium/High
Other CURL libraries	Associate Partner	Medium
MIMAS	Project Partner	Medium
JSTOR	Project Partner	Medium/High
HE Institutions	User group	High
FE Institutions	User group	High
Schools	User group	Medium

## 7. Risk Analysis

The following table identifies and analyses risks associated with the project. It indicates actions undertaken to prevent or manage the risk, which may involve: avoidance (A), minimisation (M), or contingency (C).

<b>Risk</b>	<b>Probability (1-5)</b>	<b>Severity (1-5)</b>	<b>Score (P x S)</b>	<b>Action to Prevent/Manage Risk</b>
<i>Staffing</i>				
1. Loss of key staff	2	5	10	(M) Use of secondments and existing staff for key positions. Project will provide opportunities for staff development.
<i>Organisational</i>				
2. Breakdown in consortium relationship	1	5	5	(M) UK partners are actively involved in the Project Management Group and kept informed through email list and regular updates.
3. Cost overrun	2	3	6	(M) Review deliverables and budgets.
4. Schedule overrun	3	2	6	(M) There is some contingency time within schedule.
<i>Technical</i>				
5. Equipment failure	1	2	2	(C)(M) There is spare equipment within BOPCRIS and maintenance contracts in place.
6. Software failure	3	3	9	(C)(M) Most of software is tried and tested; provision is made for maintenance of software; a programmer is on staff and further support is available.
7. Quality issues with images	3	4	12	(M) Ensure appropriate QA strategy is in place.
8. Quality issues with OCR text	3	4	12	(M) Ensure appropriate QA strategy is in place.
9. Quality issues with metadata	2	4	8	(M) Ensure appropriate QA strategy is in place.
10. Loss of data	1	5	5	(A) Ensure appropriate back-up of datasets.
<i>External suppliers</i>				
11. Quality issues with delivery collection	1	5	5	(M) Ensure appropriate QA strategy is in place.
<i>Legal</i>				
12. Copyright issues	1	5	5	(A)(C) Copyright strategy; indemnification.
13. Licensing issues	4	5	20	(M) Involvement of JISC Collections and the JISC's lawyer in the licensing drafting and negotiations.

## 8. Standards

The following table lists the standards being used in the project.

Name of standard or specification	Version	Notes
<i>Capture guidelines</i>		
JSTOR Scanning Guidelines	4.0 (January 2007), as amended for pamphlets project	These set out specifications for image capture, optimisation, and file formats.
BOPCRIS Technical Guidelines		Guidelines covering capture, OCR, and metadata, updated to reflect JSTOR's Scanning Guidelines (above)
National Preservation Office (NPO) <i>Good handling principles and practice for library and archive materials.</i> <sup>21</sup>		Guidelines covering the handling of archival materials.
<i>Archival image format</i>		
TIFF	6.0	<i>Plain text</i> is captured without scaling (100%) at 600spi bitonal (1-bit) and saved as TIFF 6.0 baseline without compression. <i>Grey or colour material</i> (e.g. illustrations or annotations) is captured without scaling (100%) at 300spi grey (8-bit) or colour (24-bit) and saved as TIFF 6.0 baseline without compression. These conform to recommendations by TASI for master images. <sup>22</sup>
<i>Archival metadata standards</i>		
METS – Metadata Encoding & Transmission Standard (structural/packaging metadata) <sup>23</sup>	1.6	METS is an established format for structuring complex digital resources such as publications. It is often used with MODS and will support the use of MIX and PREMIS. The project will aim to be <i>at least</i> minimally compliant with this standard.

<sup>21</sup> See at <http://www.bl.uk/services/npo/pdf/handling.pdf>

<sup>22</sup> "Choosing a File Format", available at: <http://www.tasi.ac.uk/advice/creating/format.html>

[Checked 3 May 2007]

<sup>23</sup> See at <http://www.loc.gov/standards/mets/>

MODS – Metadata Object Description Schema (descriptive metadata) <sup>24</sup>	3.2	MODS records are provided by MIMAS and incorporated into METS by BOPCRIS. The MODS records are at least minimally compliant with the standard.
MIX – NISO Metadata for Images in XML Schema (technical metadata) <sup>25</sup>	1.0	MIX is a fairly new standard providing XML encoding for ANSI/NISO Z39.87-2006)Z39.87: Data Dictionary - Technical Metadata for Digital Still Images. The project will aim to be <i>at least</i> minimally compliant with this standard.
PREMIS – Preservation Metadata Implementation Strategies Working Group (preservation metadata) <sup>26</sup>	1.1	PREMIS is to be used selectively within METS to support preservation of the digital resources.
<i>Archival OCR formats</i>		
TXT		BOPCRIS will generate standard ASCII text (.txt) files for each page.
IDX		BOPCRIS will additionally generate IDX files (the proprietary format used by Abbyy which includes word coordinates for use in term-highlighting
<i>Delivery image and text formats</i>		
GIF		760 pixel wide GIF images are used to display pages with plain text.
JPEG		760 pixel wide JPEG images are used to display pages with grey or colour images.
TIFF	6.0	Large multi-page TIFF images are used to support accessibility.
PDF	7.0	PDF is used to provide downloadable versions of pamphlets. These include embedded GIF and JPEG images with text in the background
<i>Delivery metadata format</i>		
NLM – National Library of Medicine Journal Archiving and Interchange Tag Set <sup>27</sup>	2.1	The archival metadata will be transformed into a metadata set based on the NLM standard.

<sup>24</sup> See at: <http://www.loc.gov/standards/mods/>

<sup>25</sup> See at <http://www.loc.gov/standards/mix/>

<sup>26</sup> See at <http://www.loc.gov/standards/premis/>

<sup>27</sup> see at <http://dtd.nlm.nih.gov/>

## 9. Technical Development

The project utilises a workflow and software that were established in the previous digitisation project. These may be refined or replaced in the course of this current project to achieve further efficiency and enhance output. Some additional software will be commissioned or acquired as outlined below.

- **Library Partners' Database.** This is a simple database commissioned from, and hosted by, the ILRT at the University of Bristol to support the preparations of the library partners. It provides a browser-based interface using PHP scripting on top of a MySQL database populated with bibliographic records received from MIMAS. It enables libraries to identify duplicates and to record: (a) items being sent; (b) their copyright status; and (c) their condition. It is also used by the Project Officer at BOPCRIS to record the arrival and departure of pamphlets at Southampton. The database was commissioned from Bristol because it needed to be in place at the beginning of the project and prior to the recruitment of the Software Developer at BOPCRIS.
- **BOPCRIS Workflow Database.** This database will be developed by the Software Developer at BOPCRIS to manage the complete workflow from the receipt of the pamphlets, through all stages of the production of the digital objects, their QA, OCR, metadata generation, delivery to JSTOR and QA by JSTOR.
- **XML Generator.** A commercial XML Editor will be used for this project to generate and edit the XML metadata. This software does not natively support METS encoded XML. Alternative open source editors that do use native METS XML will be examined and brought into use through the life of the project.
- **OCR workflow.** Adjustments to the OCR workflow have been made to enable the employment of specialist software for foreign language and old English fonts found within the pamphlets.
- **Adjustments to JSTOR's systems.** JSTOR will undertake adjustments to their systems to support the creation of the delivery dataset and the delivery to end-users.
- **Adjustments to MIMAS's systems.** MIMAS will develop software required to add the unique identifiers to the Copac records and to enable the detection of possible duplicates across library collections.
- **Development of a learning resource for JORUM.** This work will be undertaken at BOPCRIS by a Research Officer, with input from the Project Manager and Technical Project Manager and technical assistance from the Software Developer. The specific technologies are yet to be determined and will depend upon the nature of the resource.
- **Re-development of the *Pamphlets Guide*.** This development will employ standard web technologies HTML or XHTML with style sheets. The Project Team will work with MIMAS to enable querying of the Copac database by form or search string.

## 10. Intellectual Property Rights

This section describes the project's approach to IPR and licensing.

### 10.1 Copyright

Copyright is a potential IPR issue in digitising 19<sup>th</sup> century pamphlets. Although the physical collections are owned by their libraries (or included with the owners' permission), this does not automatically give libraries a right to copy the pamphlets. Most of the pamphlets will now be out of copyright and so there will be no barrier to copying. However a small proportion will be in copyright and require either permission or exclusion. The project will place each pamphlet into one of three classes: (1) Out of Copyright, (2) Uncertain Copyright, and (3) In Copyright. The table below indicates the basis for assignment of copyright status and the actions the project will take in each case.

Class	Basis	Action
1. Out of Copyright	<i>Either:</i> An early publication date, Anonymous authorship, or author's death occurred more than 70 years ago	<b>Include</b> within the collection
2. Uncertain Copyright	Author's death date is unknown	<b>Include</b> , adopting a 'notice and take-down' approach, and providing indemnification against action.
3. In Copyright	Author's death date is within past 70 years	<b>Exclude</b> , unless the partner library clears copyright.

Sampling from four of the complete collections<sup>28</sup> has found the following proportions:

Collection	Out of Copyright	Uncertain Copyright	In Copyright
Durham	76%	23%	<1%
Liverpool	83%	17%	0%
Newcastle	83%	17%	0%
UCL	100%	0%	0%

This provides support for the approach suggested above: excluding in-copyright works would have minimal impact on the quality of the selection, but excluding works of uncertain copyright status would have a detrimental effect.

The *Scoping Study* presented a workflow for determining copyright status (pp.24-26). A revised workflow is presented in Appendix D below. This was developed to enable libraries to more quickly arrive at a copyright decision when assigning status in the Library Partners' Database.

<sup>28</sup> Data was not available for Manchester's collection

## **10.2 Licensing**

[Details of licensing agreements have been excluded]

## **Project Resources**

### **11. Project Partners**

The project bid made a distinction between Primary Partners and Associate Partners. Primary Partners are those involved in Phase 1 of the pamphlets project (i.e. this current project). Associate Partners are CURL libraries who would be willing to contribute content to further phases. Associates will be kept informed of progress and may be consulted on some aspects of the current project.

The table below lists the primary project partners, their roles and main contacts

<b>Partner</b>	<b>Role</b>	<b>Contacts</b>
<b>JISC</b>	<b>Major funder</b>	<b>Alastair Dunning</b> , Programme Manager – [Email address excluded]
<b>CURL</b>	<b>Project sponsor and funder</b>	<b>Robin Green</b> , Executive Director – [Email address excluded]
<b>University of Southampton/ BOPCRIS</b>	<b>Project lead and digitisation partner</b>	<b>Mark Brown</b> , University Librarian and Project Director – [Email address excluded] <b>Grant Young</b> , Project Manager (secondment from Bristol) – [Email address excluded] <b>Julian Ball</b> , BOPCRIS Manager and Technical Project Manager – [Email address excluded] <b>Richard Wake</b> , Deputy Librarian – [Email address excluded]
<b>JSTOR</b>	<b>Business partner – responsible for preservation and delivery</b>	<b>Michael Spinella</b> , Executive Director – [Email address excluded] <b>Kimberly Lutz</b> , Director of Publisher Relations – [Email address excluded] <b>John Kiplinger</b> , Director of Production – [Email address excluded]

<b>MIMAS</b>	<b>Bibliographic services partner</b> – responsible for supplying catalogue records and links to digital versions	<b>Sean Dunne</b> , Bibliographic & Scientific Services Manager – [Email address excluded]
<b>University of Bristol</b>	<b>Library partner</b>	<b>Peter King</b> , Director of Library Services and Chair of Project Steering Group; <b>David Wilkins</b> , Cataloguing – [Email address excluded]
<b>Durham University</b>	<b>Library partner</b>	<b>Sheila Hingley</b> , Head of Heritage Collections – [Email address excluded]
<b>University of Liverpool</b>	<b>Library partner</b>	<b>Katy Hooper</b> , Special Collections Librarian – [Email address excluded]
<b>LSE</b>	<b>Library partner</b>	<b>Barbara Humphries</b> , Assistant Librarian, Information Services Department – [Email address excluded]
<b>University of Manchester</b>	<b>Library partner</b>	<b>Jenny Curtis</b> , Head of Digital Programmes – [Email address excluded]
<b>University of Newcastle</b>	<b>Library partner</b>	<b>Melanie Wood</b> , Special Collections and Archives Librarian – [Email address excluded]
<b>UCL</b>	<b>Library partner</b>	<b>Lesley Pitman</b> , Librarian and Director of Information Services, School of Slavonic and East European Studies Library – [Email address excluded]

The following are Associate Partners:

British Library  
National Library of Scotland  
National Library of Wales  
Trinity College Dublin  
King's College London  
University of Aberdeen  
University of Birmingham  
University of Cambridge  
University of Cardiff  
University of Edinburgh

University of Glasgow  
University of Leeds  
University of London  
University of Nottingham  
University of Oxford  
University of Reading  
University of Sheffield  
University of Warwick  
V&A

[Discussion of licensing agreements has been excluded]

## 12. Project Management

The project is managed by a core Project Team, comprised of the Project Director, Project Manager, Technical Project Manager, CURL Executive Director (representing the interests of CURL), and Southampton Deputy Librarian (overseeing the budget for Southampton). There is also a team of staff within Southampton's BOPCRIS unit responsible for creating the digital content, and staff within partner organisations (the libraries, MIMAS, and JSTOR) undertaking other work. Details of the Project Team and the BOPCRIS Team are provided in sections 12.1 and 12.2 below.

Two groups support the Project Team: the Project Management Group (PMG), which includes representation from each UK partner, and the Project Steering Group (PSG), which includes academic, library and digitisation representatives. The membership and roles of these groups is described in sections 12.3 and 12.4 below.

### 12.1 Project Team

Role	Roleholder	Allocation of time/budget	Responsibilities and reporting
Project Director	Mark Brown	As required/ not funded from project budget	Overall responsibility for the project and its budget, supervision of Project Manager and Technical Project Manager
Project Manager	Grant Young	0.5	Overall project management and monitoring, risk management, reporting, liaison with partners, and project dissemination. Reports to the Project Director.
Technical Project Manager	Julian Ball	0.5	Management of the BOPCRIS production processes and staff, quality assurance, liaison with partners over technical issues. Reports to the Project Director.
Budget Manager	Richard Wake	As required/ not funded from project budget	Management of the budget for the project. Reports to the Project Director.
CURL Representative	Robin Green	As required/ not funded from project budget	Represents CURL in decision-making, particularly in regard to licensing.

## 12.2 BOPCRIS Team

Role	Roleholder	Allocation of time/budget	Responsibilities
BOPCRIS Manager	Julian Ball	As above	As above
Inventory Project Officer	Maureen Langham	1.0	Oversees the transportation of pamphlets between the libraries and BOPCRIS and the flow of pamphlets through the BOPCRIS production system. Reports to the Technical Project Manager.
QA Project Officer	Verna Acres	1.0	Oversees the metadata creation and QA processes within BOPCRIS. Reports to the Technical Project Manager.
Software Developer	Ed Fay	0.6	Undertakes development of the BOPCRIS production system required to manage pamphlet production and generate required metadata. Reports to the Technical Project Manager
Research Officer	TBA	1.0 for 6 month in 2008	Works with BOPCRIS staff and the Project Manager to re-develop the online <i>Guide</i> and create a learning resource for JORUM. Reports to the Technical Project Manager.
Scanning operators	TBA	5.0	Undertake digitisation and quality assurance work. Report to the Project Officers and Technical Project Manager.

## 12.3 Project Management Group

The Project Management Group (PMG) includes representation from all UK partners. It operates in a similar way to a Project Board within the PRINCE2 methodology, overseeing and guiding the management of the project and dealing with any exceptional circumstances. The PMG has regular meetings (3-4 times a year) and will meet at other times if required. In addition to formal meetings, the group has a private email forum on JISMAIL (PAMPHLET-DIGI).

The Project Management Group includes:

Mark Brown (Project Director) (Chair)  
 Grant Young (Project Manager)  
 Julian Ball (Technical Project Manager)  
 Robin Green (CURL)  
 Alastair Dunning (JISC)  
 Sean Dunne (MIMAS)

David Wilkins (Bristol)  
 Sheila Hingley (Durham)  
 Katy Hooper (Liverpool)  
 Barbara Humphries (LSE)  
 Jenny Curtis (Manchester)  
 Melanie Wood (Newcastle)  
 Lesley Pitman (UCL)

### **12.4 Project Steering Group**

The Project Steering Group (PSG) provides strategic oversight. With membership drawn from research, teaching, library and digitisation communities, it is tasked with both scrutinising and supporting the project. The PSG meets at the beginning, middle and near the end of the project although members may be consulted at other times.

The Project Steering Group includes:

Dr Peter King (Chair)	Librarian, University of Bristol
Prof John Belchem	Historian, University of Liverpool
Alastair Dunning	Programme Manager, JISC
Robin Green	Executive Director, CURL
Dr Tim Leunig	Economic Historian, LSE
Ronald Milne	Director of Scholarship and Collections, British Library
Christine Paterson	History Teacher, West Kirby Grammar School, Wirral
Dr James Thompson	Historian, University of Bristol
Dr Matthew Woollard	Head of Digital Preservation and Systems, UK Data Archive, University of Essex

### **13. Programme Support**

As described in section 11 (above), the project will require some additional support in negotiating and concluding the licensing agreements and in paying for this work.

The project may also require further support and funding in securing a UK repository for the archival dataset (see 3.2.5 above) as this was not required or budgeted for when the revised bid was prepared.

### **14. Budget**

Please see Appendix A for the budget.

[Discussion of budget has been excluded]

## Detailed Project Planning

### 15. Workpackages

Please see Appendix B for a description of the workpackages.

### 16. Evaluation Plan

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Ongoing from beginning of digitisation	<b>Archival dataset</b> (images, OCR-text and metadata)	Has project met target of 1 million pages? Is data of sufficient quality? Is data delivered on time?	Page image tally. Quality plan (below) and schedule. Formal independent evaluation at end of project	1 million pages achieved. Quality targets achieved.
Ongoing from first delivery of data to JSTOR	<b>Delivery dataset</b> (images, OCR-text and metadata)	Is data of sufficient quality? Is data delivered on time?	Quality plan (below) and schedule. Survey of users. Formal independent evaluation at end of project	90% of reviewers/users are happy with <i>quality</i> of online pamphlets.
Ongoing from first batch of content publicly released (Spring or Summer 2008)	<b>Online pamphlet collection</b>	Does <i>range of content</i> of collection satisfy users? Does <i>presentation</i> of collection satisfy users? Is collection usable and accessible?	Review by peers and Steering Group. Survey of users. Formal independent evaluation at end of project	90% of reviewers/users are happy with <i>selection</i> of pamphlets. 90% of reviewers/users are happy with <i>presentation</i> of pamphlets. Collection conforms to JSTOR's accessibility and usability standards
Ongoing from release of records in Copac	<b>Enhanced catalogue records</b>	Do links to the pamphlets work in Copac and local catalogues?	Spot checks.	100% of pamphlets checked can be accessed

Year 2 of project	<b>Redeveloped Pamphlets Guide</b>	Does Guide assist users in understanding pamphlet literature and searching by theme or collection?	User feedback (via email address). Review by peers and Steering Group. Analysis of logs (if available).	90% of feedback from reviewers/users is positive. 50% increase in usage of Guide within six months of completion.
Year 2 of project	<b>Learning resource for JORUM</b>	Do users find the resource useful? Does it provide a good model for other resources?	User feedback (via email address). Review by peers and Steering Group. Statistics (if available) from JORUM	90% of feedback from reviewers/users is positive. Statistics (if available) show steady growth in usage.
Summer 2008	<b>Digitisation resource 'toolkit'</b>	Is information useful to others?	Peer review	
Summer 2008	<b>Day conference</b>	Does the event meet set objectives?	Formal evaluation at end of the event	80% of attendees rate the event overall as good or excellent
Ad hoc	<b>Other Presentations and conferences</b>	Do the presentations promote the collection and contribute to knowledge?	Peer review	
Ad hoc	<b>Articles in publications</b>	Do the presentations promote the collection and contribute to knowledge?	Peer review	
Ongoing	<b>Project Documentation</b>	Is documentation clear? Does it meet needs and requirements?	Feedback from partners and JISC	
mid-October 2008-mid-January 2009	<b>Overall success of project</b>	How well did the project meet its aim and objectives?	Feedback from users. Feedback from JISC. Formal independent evaluation at end of project	Success in above measures. External reviewer regards project as "overall a success."

## 17. Quality Plan

<b>Output: Archival Images</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Ongoing	Fidelity to original; legibility (all detail captured)	Every image is checked by two staff at BOPCRIS. Further checking by JSTOR (approx. 10% of images).		BOPCRIS; JSTOR	

<b>Output: OCR text</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Ongoing	Accuracy of text	Spot checking by BOPCRIS. Further checking by JSTOR (approx. 10% of files)		BOPCRIS; JSTOR	Manual count initially; moving to open source tool

<b>Output: Metadata</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Ongoing	Compliant with formal standards.	XML validation against Library of Congress schemas.	Validation report	BOPCRIS	XML editor

<b>Output: Delivery data</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Ongoing	User satisfaction	User survey	90% of feedback is positive	Project Team and JSTOR	Survey

<b>Output: Enhanced Copac records</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Ongoing	Links accurate	Spot checking	100% accuracy	JSTOR, MIMAS, BOPCRIS	

<b>Output: Redeveloped Pamphlets Guide</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Once complete, in Year 2 of project	User satisfaction	User feedback (via email)	90% of feedback is positive	Project Team	

<b>Output: Learning resource for JORUM</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)

Once complete, in Year 2 of project	User feedback	User feedback (email)	90% of feedback is positive	Project Team	
-------------------------------------	---------------	-----------------------	-----------------------------	--------------	--

<b>Output: Digitisation “toolkit”</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Once complete in 2008	User feedback	User feedback (email)	90% of feedback is positive	Project Team	

<b>Output: Day conference</b>					
Timing	Quality criteria	QA method(s)	Evidence of compliance	Quality responsibility	Quality tools (if applicable)
Once complete in 2008	Attendee evaluation	Formal evaluation	80% of attendees rate the event as overall good or excellent	Project Team	

## 18. Dissemination Plan

Activity	Purpose	Target audience	Timing	Key message
<b>Press releases</b>	Raise awareness	General	At beginning and at key milestones	Project is underway and succeeding.
<b>Public Webpage</b> (on CURL site)	Raise awareness; provide key information	All interested parties	Established at beginning and updated regularly	General project news.
<b>Entry in Intute</b>	Raise awareness	Subject specialists	By middle of year 1, once sufficient content on Public Website. Updated when collection becomes available.	Project is underway and worth keeping in touch with.
<b>Public JISMAIL list</b> (PAMPHLETS-ONLINE)	Enable interested parties to be informed of project news.	All interested parties	Established at beginning. Regular postings.	General project news
<b>Private JISMAIL list</b> (PAMPHLET-DIGI)	Forum for Project Management Group (PMG); keep partners informed; secure space for project documents.	Project Management Group	Established at beginning. Regular postings.	Detailed project news and discussion.

<b>Private JISCMail list</b> (JISC-METADATA)	Share experience with XML metadata standards.	JISC and wider community.	Established at beginning. Regular postings.	Exchanging experience.
<b>Postings to relevant mailing lists and blogs.</b>	Raise awareness of project and its publications/events	Specialist audiences (subject-related, e-learning related, digitisation-related)	Throughout life of project	Project is underway and providing useful deliverables.
<b>Day conference</b>	Raise awareness, launch resource, disseminate learning	Specialists (research, teaching or digitisation - TBC)	Summer 2008	Pamphlet collection is a useful resource; project has uncovered useful lessons.
<b>“Toolkit”</b>	Disseminate learning.	Digitisation specialists	Summer 2008	Project has uncovered useful lessons.
<b>Learning resource</b>	Demonstrate how collection can be repurposed for teaching and learning application.	Teachers and learners; resource developers.	Year 2 of project	Pamphlets can be useful in teaching and learning.
<b>Pamphlets Guide</b>	Provide focus for collections. Encourage use.	Specialist and non-specialist audiences	Year 2 of project	Pamphlets can be useful for research and teaching and learning.
<b>Participation in JISC events and publications</b>	Raise awareness, update on progress, and contribute to knowledge.	JISC’s audiences	As required and appropriate	Progress of project, dissemination of learning.
<b>Other conference contributions</b> (to be advised)	Raise awareness, update on progress, and contribute to knowledge.	Specialists (research, teaching or digitisation)	TBC	Progress of project, dissemination of learning.
<b>Articles in other publications</b> (to be advised)	Raise awareness, update on progress, and contribute to knowledge.	Specialists (research, teaching or digitisation)	TBC	Progress of project, dissemination of learning.
<b>Case Studies</b> (to be advised, possibly with TASI or AHDS)	Raise awareness, disseminate learning.	Digitisation specialists	Towards end of project	Pamphlet has uncovered useful lessons.

## 19. Exit and Sustainability Plans

### 19.1 Exit Plan

Project outputs	Action for Take-up & Embedding	Action for Exit
<b>Archival dataset</b> (images, OCR-text and metadata)	Not available for general use	Use of archival dataset is governed by CURL-JSTOR agreement. Contributing libraries entitled to hold and use copies.
<b>Delivery dataset</b> (images, OCR-text and metadata)	See Dissemination plan (section 18)	Use of delivery dataset is governed by CURL-JSTOR agreement.
<b>Online pamphlet collection</b>	See Dissemination plan (section 18)	Use of delivery dataset is governed by CURL-JSTOR agreement.
<b>Enhanced catalogue records</b>	MIMAS to make available to CURL libraries	Records are held by MIMAS
<b>Redeveloped Pamphlets Guide</b>	See Dissemination plan (section 18)	Guide is hosted by CURL
<b>Learning resource for JORUM</b>	See Dissemination plan (section 18)	Resource is held within JORUM
<b>Digitisation resource 'toolkit'</b>	See Dissemination plan (section 18)	Resource is hosted on Project Website or by JISC service (e.g. TASI)
<b>Day Conference</b>	See Dissemination plan (section 18)	Presentations hosted on Project Website; May be incorporated into 'toolkit' if event has digitisation focus.
<b>Presentations at other conferences/events</b>	See Dissemination plan (section 18)	Presentations made available by conferences and held on Project Website.
<b>Articles in publications</b>	See Dissemination plan (section 18)	Articles published and (where possible) held on Project Website and within appropriate repositories.
<b>Project Documentation</b>	Some of documentation made available to public via project website and other deliverables. Some kept confidential.	Held by lead institution and (in part and as required) by JISC.
<b>Scholarship into the pamphlet form</b>	Promote scholarship in other deliverables and dissemination.	
<b>Increased use in research and teaching</b>	Promote use of collection in other deliverables and dissemination.	
<b>Methodology lessons</b>	Will be embedded within BOPCRIS and within deliverables ('toolkit', presentations, articles)	BOPCRIS will hold this knowledge; also captured in some of other deliverables.
<b>Metadata lessons</b>	Will be embedded within BOPCRIS and within deliverables ('toolkit', presentations, articles)	BOPCRIS will hold this knowledge; also captured in some of other deliverables.

## 19.2 Sustainability Plan

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
<b>Archival dataset</b> (images, OCR-text and metadata)	Sustained through business model (agreement with JSTOR)	Periodic review of agreement with JSTOR. Alternative delivery possible. Copies held in UK by libraries and additional repository for FE/HE community.	Ensure CURL-JSTOR licensing agreement provides security and flexibility.
<b>Delivery dataset</b> (images, OCR-text and metadata)	Sustained through business model (agreement with JSTOR)	Periodic review of agreement with JSTOR. Alternative delivery possible. Back-up of delivery dataset held by MIMAS.	Ensure CURL-JSTOR licensing agreement provides security and flexibility.
<b>Online pamphlet collection</b>	Sustained through business model (agreement with JSTOR)	Periodic review of agreement with JSTOR. Alternative delivery possible.	Ensure licensing agreement provides security and flexibility.
<b>Enhanced catalogue records</b>	Maintained by MIMAS	Extended through further phases/projects	None anticipated.
<b>Redeveloped Pamphlets Guide</b>	Hosted by CURL	Developed further through later phases/projects.	Unless taken forward through further projects will require some level of checking and maintenance.
<b>Learning resource for JORUM</b>	Hosted by JORUM	Maintained by JORUM	None anticipated.
<b>Digitisation resource 'toolkit'</b>	Hosted by JISC service	Maintained by JISC service	Agreement with host over maintenance
<b>Day Conference</b>	Deliverables on Project Website	Project Website maintained by CURL	Project Website will require some level of checking and maintenance.
<b>Presentations at conferences/events</b>	Held on Project Website and other sites	Presentations held in multiple locations.	None anticipated.
<b>Articles in publications</b>	Held on Project Website, publications, repositories.	Articles held in multiple locations.	None anticipated.
<b>Project Documentation</b>	Held by lead institution and JISC	Maintained by lead institution and JISC	Documentation will require some level of checking and maintenance.
<b>Scholarship into pamphlet form</b>	Pamphlets will continue to be available	Beyond control of project or partners.	Beyond control of project or partners.
<b>Increased use in research and</b>	Anticipated increase in citations and links	Beyond control of project or partners.	Beyond control of project or partners.

<b>teaching</b>	to pamphlets, incorporation into learning resources.		
<b>Methodology learnings</b>	Held by BOPCRIS and captured in deliverables.	Incorporated into future BOPCRIS work.	None anticipated.
<b>Metadata learnings</b>	Held by BOPCRIS and captured in deliverables.	Incorporated into future BOPCRIS work.	None anticipated.

## ***Appendixes***

### **Appendix A. Project Budget**

[Details of the budget have been excluded. The total grant awarded to this project by JSC is £778,991]

[Details of the budget have been excluded. The total grant awarded to this project by JSC is £778,991]

## Appendix B. Workpackages

<b>WORKPACKAGES</b>	<b>Month</b>	<b>1 MAR-07</b>	<b>2 APR-07</b>	<b>3 MAY-07</b>	<b>4 JUN-07</b>	<b>5 JUL-07</b>	<b>6 AUG-07</b>	<b>7 SEP-07</b>	<b>8 OCT-07</b>	<b>9 NOV-07</b>	<b>10 DEC-07</b>	<b>11 JAN-08</b>	<b>12 FEB-08</b>
1: Pre-Project activities		X	X										
2: Project Management		X	X	X	X	X	X	X	X	X	X	X	X
3: Development		X	X	X	X	X	X	X	X	X	X	X	X
4: Selection & Preparation			X	X	X	X	X	X	X	X	X	X	X
5: Production				X	X	X	X	X	X	X	X	X	X
6: Delivery					X	X	X	X	X	X	X	X	X
7: Preservation				X	X	X	X	X	X	X	X	X	X
8: Linking							X	X	X	X	X	X	X
9: Marketing & Dissemination		X	X	X	X	X	X	X	X	X	X	X	X
10: Evaluation													

<b>WORKPACKAGES</b>	<b>Month</b>	<b>13 MAR-08</b>	<b>14 APR-08</b>	<b>15 MAY-08</b>	<b>16 JUN-08</b>	<b>17 JUL-08</b>	<b>18 AUG-08</b>	<b>19 SEP-08</b>	<b>20 OCT-08</b>	<b>21 NOV-08</b>	<b>22 DEC-08</b>	<b>23 MAR-09</b>	<b>24 FEB-09</b>
1: Pre-Project activities													
2: Project Management		X	X	X	X	X	X	X	X	X	X	X	X
3: Development													
4: Selection & Preparation		X	X	X	X	X	X	X	X	X	X	X	X
5: Production		X	X	X	X	X	X	X	X	X	X	X	X
6: Delivery		X	X	X	X	X	X	X	X	X	X	X	X
7: Preservation		X	X	X	X	X	X	X	X	X	X	X	X
8: Linking		X	X	X	X	X	X	X	X	X	X	X	X
9: Marketing & Dissemination		X	X	X	X	X	X	X	X	X	X	X	X
10: Evaluation			X	X	X	X	X	X	X	X	X	X	

**Project start date:** 01-03-2007

**Project completion date:** 28-02-2009

**Duration:** 24 months

See also Gantt charts below

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables & reports in bold)	Milestone	Responsibility (see below for key)
<b>WORKPACKAGE 1: Pre-Project and initiation activities</b>  <b>Objective: To enable the project to begin promptly in March 2007.</b>					
Secondment of Project Manager	01/01/07	01/03/07	Project Manager in post	<b>Yes</b>	GY/MB/PK
Recruitment of Software Developer, Project Officers	01/01/07	01/03/07	Software Developer and Project Officer in post	<b>Yes</b>	JB
Establishment of Project Management Group (PMG)	01/01/07	01/03/07	Project Management Group constituted	<b>Yes</b>	MB
Establishment of Project Steering Group (PSG)	01/01/07	31/05/07	Project Steering Group constituted	<b>Yes</b>	PK
Finalisation of Pamphlet preparation workflow	01/01/07	01/04/07	<b>Pamphlet preparation strategy/workflow</b> (see appendix C)	<b>Yes</b>	GY
Finalisation of Copyright strategy and workflow	01/01/07	01/04/07	<b>Copyright strategy/workflow</b> (see appendix D)	<b>Yes</b>	GY
Finalisation of collection schedule	01/01/07	01/03/07	<b>Revised pamphlets schedule</b> (see Workpackage 4 and Gantt chart below)	<b>Yes</b>	GY

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables & reports in bold)	Milestone	Responsibility (see below for key)
Acquisition of MODS records from MIMAS	01/01/07	28/02/07	MODS records received	Yes	GY
<b>WORKPACKAGE 2: Project Management activities</b>					
<i><b>Objective: To ensure that all the Work Packages are managed coherently and that all the project outputs are delivered within agreed deadlines and budgets.</b></i>					
Preparation of Project Plan	01/03/07	31/05/07	<b>Project Plan</b> completed and delivered to JISC Programme Manager	Yes	GY/ RW(budget)
1 <sup>st</sup> Progress Report (30/09/07)	01/09/07	30/09/07	<b>Progress Report</b>	Yes	GY/ RW(budget)
2 <sup>nd</sup> Progress Report (31/01/08)	01/01/08	31/01/08	<b>Progress Report</b>	Yes	GY/ RW(budget)
3 <sup>rd</sup> Progress Report (31/05/08)	01/05/08	31/05/08	<b>Progress Report</b>	Yes	GY/ RW(budget)
4 <sup>th</sup> Progress Report (30/09/08)	01/09/08	30/09/08	<b>Progress Report</b>	Yes	GY/ RW(budget)
5 <sup>th</sup> Progress Report (31/02/09)	01/01/09	31/01/09	<b>Progress Report</b>	Yes	GY/ RW(budget)
Final Report (draft)	01/12/08	31/01/09	Draft Final Report	Yes	GY/ RW(budget)
Final Report	01/12/08	28/02/09	<b>Final Report</b>	Yes	GY/ RW(budget)
Completion Report	01/12/08	28/02/09	<b>Completion Report</b>	Yes	GY/ RW(budget)
Project web page on JISC website	01/03/07	31/03/07	<b>Web page on JISC site</b>	Yes	JISC
Project website (on CURL page)	01/03/07	31/05/07	<b>Project website</b> is live	Yes	GY/RG

<b>Workpackage and activity</b>	<b>Earliest start date</b>	<b>Latest completion date</b>	<b>Outputs</b> (deliverables & reports in bold)	<b>Milestone</b>	<b>Responsibility</b> (see below for key)
Consortium Agreement	01/03/07	31/05/07	<b>Consortium agreement</b> (agreement 3 in section 10.2 above)	<b>Yes</b>	Project Team
HEFCE licence	01/03/07	28/02/08	<b>HEFCE licence</b> (agreement 9 in section 10.2 above)	<b>Yes</b>	Project Team
JORUM licence	01/03/07	28/02/08	<b>JORUM licence</b> (agreement 10 in section 10.2 above)	<b>Yes</b>	Project Team
JISC Collections Agreement	01/03/07	31/05/07	<b>JISC Collections agreement</b> (agreement 5 in section 10.2 above)	<b>Yes</b>	Project Team
Howick and FCO agreements	01/03/07	31/12/07	<b>Howick and FCO agreements</b> (agreements 1 and 2 in section 10.2 above)	<b>Yes</b>	Project Team
CURL-JSTOR	01/03/07	31/05/07	<b>CURL-JSTOR agreement</b> (agreement 4 in section 10.2 above)	<b>Yes</b>	Project Team
JSTOR Collection Sub-licence	01/03/07	28/02/08	<b>JISC Collection sub-licence</b> (agreement 6 in section 10.2 above)	<b>Yes</b>	Project Team
MOU between Southampton and Liverpool	01/03/07	31/04/07	<b>MOU Southampton-Liverpool</b>	<b>Yes</b>	JB
MOU between Southampton and Newcastle	01/03/07	31/05/07	<b>MOU Southampton-Newcastle</b>	<b>Yes</b>	JB
MOU between Southampton and Bristol	01/03/07	30/06/07	<b>MOU Southampton-Bristol</b>	<b>Yes</b>	JB
MOU between Southampton and UCL	01/03/07	30/09/07	<b>MOU Southampton-UCL</b>	<b>Yes</b>	JB
MOU between Southampton and Durham	01/03/07	31/12/07	<b>MOU Southampton-Durham</b>	<b>Yes</b>	JB
MOU between Southampton and Manchester	01/03/07	28/02/08	<b>MOU Southampton-Manchester</b>	<b>Yes</b>	JB
MOU between Southampton and LSE	01/03/07	15/05/08	<b>MOU Southampton-LSE</b>	<b>Yes</b>	JB
Feb Management Group Meeting	13/02/07	13/02/07	PMG Meeting at UCL	<b>Yes</b>	MB/GY
May Management Group Meeting	15/05/07	15/05/07	PMG Meeting at Southampton	<b>Yes</b>	MB/GY
Nov Management Group Meeting	13/11/07	13/11/07	PMG Meeting (venue TBC)	<b>Yes</b>	MB/GY
Feb Management Group Meeting	12/02/08	12/02/08	PMG Meeting at LSE	<b>Yes</b>	MB/GY
May Management Group Meeting	13/05/08	13/05/08	PMG Meeting (venue TBC)	<b>Yes</b>	MB/GY
Nov Management Group Meeting	TBC	TBC	PMG Meeting (venue TBC)	<b>Yes</b>	MB/GY
Feb Management Group Meeting	TBC	TBC	PMG Meeting (venue TBC)	<b>Yes</b>	MB/GY
1 <sup>st</sup> Project Steering Group Meeting	04/05/07	04/05/07	PSG Meeting at British Library	<b>Yes</b>	PK/GY
2 <sup>nd</sup> Project Steering Group Meeting	TBC	TBC	PSG Meeting at Southampton	<b>Yes</b>	PK/GY

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables & reports in bold)	Milestone	Responsibility (see below for key)
3 <sup>rd</sup> Project Steering Group Meeting	TBC	TBC	PSG Meeting (venue TBC)	Yes	PK/GY
<b>WORKPACKAGE 3: Development work</b>					
<i><b>Objective: To make necessary adjustments to existing systems or develop new systems to support the major work packages (i.e. WP4-WP8).</b></i>					
Development of Library Partners' Database	01/01/07	28/02/07	<b>Database</b> ready to use	Yes	GY
Development of BOPCRIS workflow database	01/05/07	31/07/07	<b>Database</b> ready to use	Yes	JB/EF
Adjustments to OCR workflow	01/01/07	30/09/08	System ready and optimised for production (long timeframe represents series of improvements)	Yes	JB/EF
Define metadata sets	01/01/07	30/09/08	All XML metadata defined	Yes	JB/EF
Customisation of XML Editor	01/01/07	30/09/08	System ready to produce metadata (long timeframe represents series of improvements)	Yes	JB/EF
Adjustments to the JSTOR system	01/06/07	28/02/08	System ready to deliver pamphlet collection	Yes	JSTOR
MIMAS systems for incorporating links	01/10/07	31/12/07	Systems in place for incorporating links in Copac and library OPACs	Yes	MIMAS
Development of JORUM resource (see WP9 below)					
Re-development of <i>Pamphlets Guide</i> (see WP9 below)					
<b>WORKPACKAGE 4: Selection, Preparation and delivery of Pamphlet collections</b>					
<i><b>Objective: To efficiently select, prepare and transport pamphlets from and back to libraries</b></i>					

<b>Workpackage and activity</b>	<b>Earliest start date</b>	<b>Latest completion date</b>	<b>Outputs</b> (deliverables & reports in bold)	<b>Milestone</b>	<b>Responsibility</b> (see below for key)
<b>Liverpool collection</b>					
Preparations	01/04/07	31/04/07	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/05/07	01/05/07	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	31/08/07	31/08/07	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>Newcastle collection</b>					
Preparations	01/05/07	31/05/07	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/06/07	01/06/07	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	30/09/07	30/09/07	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>Bristol collection (first batch)</b>					
Preparations	01/05/07	30/06/07	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/07/07	01/07/07	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	31/12/07	31/12/07	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>UCL collection</b>					
Preparations	01/08/07	30/09/07	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/10/07	01/10/07	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	31/03/08	31/03/08	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>Durham collection</b>					
Preparations	01/12/07	31/12/07	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/01/08	01/01/08	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	30/04/08	30/04/08	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>Manchester collection</b>					
Preparations	01/01/08	29/02/08	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/03/08	01/03/08	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	31/07/08	31/07/08	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>LSE collection (first batch)</b>					

<b>Workpackage and activity</b>	<b>Earliest start date</b>	<b>Latest completion date</b>	<b>Outputs</b> (deliverables & reports in bold)	<b>Milestone</b>	<b>Responsibility</b> (see below for key)
Preparations	01/03/08	15/05/08	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	16/05/08	16/05/08	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	15/09/08	15/09/08	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>Bristol collection (second batch)</b>					
Preparations	15/05/08	31/07/08	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	01/08/08	01/08/08	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	31/08/08	31/08/08	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>LSE collection (second batch)</b>					
Preparations	1/09/08	15/11/08	Pamphlets recorded on database and packed for delivery to BOPCRIS	<b>Yes</b>	LIB/GY/JB
Delivery to BOPCRIS	16/11/08	16/11/08	Pamphlets received by BOPCRIS	<b>Yes</b>	ML/JB
Return to library	15/03/09	15/03/09	Pamphlets received by library	<b>Yes</b>	ML/JB
<b>WORKPACKAGE 5: Production</b>					
<b><i>Objective: To create high-quality digital images, metadata and OCR</i></b>					
Production of archival dataset for <b>Liverpool</b> collection	01/05/07	30/06/07	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>Newcastle</b> collection	01/06/07	31/07/07	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>Bristol</b> collection (first batch)	01/07/07	15/10/07	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>UCL</b> collection	01/10/07	15/01/08	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>Durham</b> collection	01/01/08	15/03/08	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>Manchester</b> collection	01/03/08	31/05/08	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB

<b>Workpackage and activity</b>	<b>Earliest start date</b>	<b>Latest completion date</b>	<b>Outputs</b> (deliverables & reports in bold)	<b>Milestone</b>	<b>Responsibility</b> (see below for key)
Production of archival dataset for <b>LSE</b> collection (first batch)	16/05/08	15/08/08	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>Bristol</b> collection (second batch)	01/08/08	15/11/08	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
Production of archival dataset for <b>LSE</b> collection (second batch)	16/11/08	31/01/09	<b>Archival dataset (images, OCR, metadata)</b>	<b>Yes</b>	JB
<b>WORKPACKAGE 6: Delivery</b>					
<i><b>Objective: To effectively deliver the collection to users</b></i>					
Production of delivery dataset	15/06/07	31/03/09	<b>Delivery dataset</b>	<b>Ongoing</b>	JSTOR
Release of content to users	01/01/08	31/04/09	<b>Collection</b> available for use	<b>Ongoing</b>	JSTOR
<b>WORKPACKAGE 7: Preservation</b>					
<i><b>Objective: To ensure the long-term preservation (including future upgrades and migrations as technology changes) and accessibility of the material</b></i>					
Securing of UK repository for archival dataset	01/01/08	28/02/09	UK repository is arranged for the archival dataset	<b>Yes</b>	Project Team
Archiving of archival dataset by JSTOR and UK repository	01/08/07	Ongoing, beyond life of project	Data is stored, checked and migrated as appropriate	<b>Ongoing</b>	JSTOR and UK Repository
Provision of archived datasets to contributing libraries	01/01/08	Ongoing, beyond life of project	Data is provided to UK partners on request	<b>Ongoing</b>	JSTOR and UK Repository

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables & reports in bold)	Milestone	Responsibility (see below for key)
<b>WORKPACKAGE 8: Linking</b>  <i><b>Objective: To achieve linking from Copac, from collection descriptions in the Pamphlets Guide, and from the individual OPACs of libraries holding the pamphlets</b></i>					
Delivery of URLs/DOIs from JSTOR to MIMAS	01/08/07	30/04/09	URLs/DOIs available for linking	<b>Ongoing</b>	JSTOR
Provision of links for Copac, libraries and other services	01/01/08	Ongoing, beyond life of project	<b>Enhanced catalogue records</b>	<b>Ongoing</b>	MIMAS/LIB
<b>WORKPACKAGE 9: Marketing and dissemination</b>  <i><b>Objective: To generate wide interest in the project and wide usage of its collection and to publicise lessons learned</b></i>					
Development of website and publicity materials	01/01/07	28/02/09	<b>Up-to-date website and marketing materials</b> available	<b>Ongoing</b>	GY
Press releases, postings to lists, listings in web directories	01/01/07	28/02/09	<b>Press releases and entries in directories</b>	<b>Ongoing</b>	GY
Presentations and articles	01/06/07	28/02/09	<b>Presentations and articles</b>	<b>Ongoing</b>	GY/JB
Day conference	30/09/08 (tentative)	30/09/08 (tentative)	<b>Day conference</b> takes place	<b>Yes</b>	GY/Project Team
Creation of "toolkit"	01/08/08	31/01/09	<b>Toolkit</b> is made available	<b>Yes</b>	GY/JB/EF
Creation of learning resource for JORUM	01/05/08	31/10/08	Learning resource is made available in JORUM	<b>Yes</b>	RO/EF/Project Team
Redevelopment of the Pamphlets Guide	01/05/08	31/10/08	Redeveloped <b>Pamphlets Guide</b> is made available online.	<b>Yes</b>	RO/EF/Project Team

Workpackage and activity	Earliest start date	Latest completion date	Outputs (deliverables & reports in bold)	Milestone	Responsibility (see below for key)
<b>WORKPACKAGE 10: Evaluation</b>  <i><b>Objective: To ensure quality and provide external evaluation study whose assessment and recommendations will be incorporated into the Final Report</b></i>					
Survey work with users	01/04/08	30/06/08	<b>Reports</b>	<b>Ongoing</b>	GY
Summative external evaluation	01/11/08	31/01/09	<b>External evaluation report</b>	<b>Yes</b>	GY

**Responsibilities:**

**EF** = Ed Fay, Software Developer

**GY** = Grant Young, Project Manager

**JB** = Julian Ball, Technical Project Manager

**JSTOR**

**LIB** = Primary Partner Libraries

**MB** = Mark Brown, Project Director

**MIMAS**

**ML** = Maureen Langham, Inventory Project Officer

**Project Team** – see 12.1 above

**PK** = Peter King, Chair of Steering Group

**RG** = Robin Green, CURL representative

**RO** = Research Officer

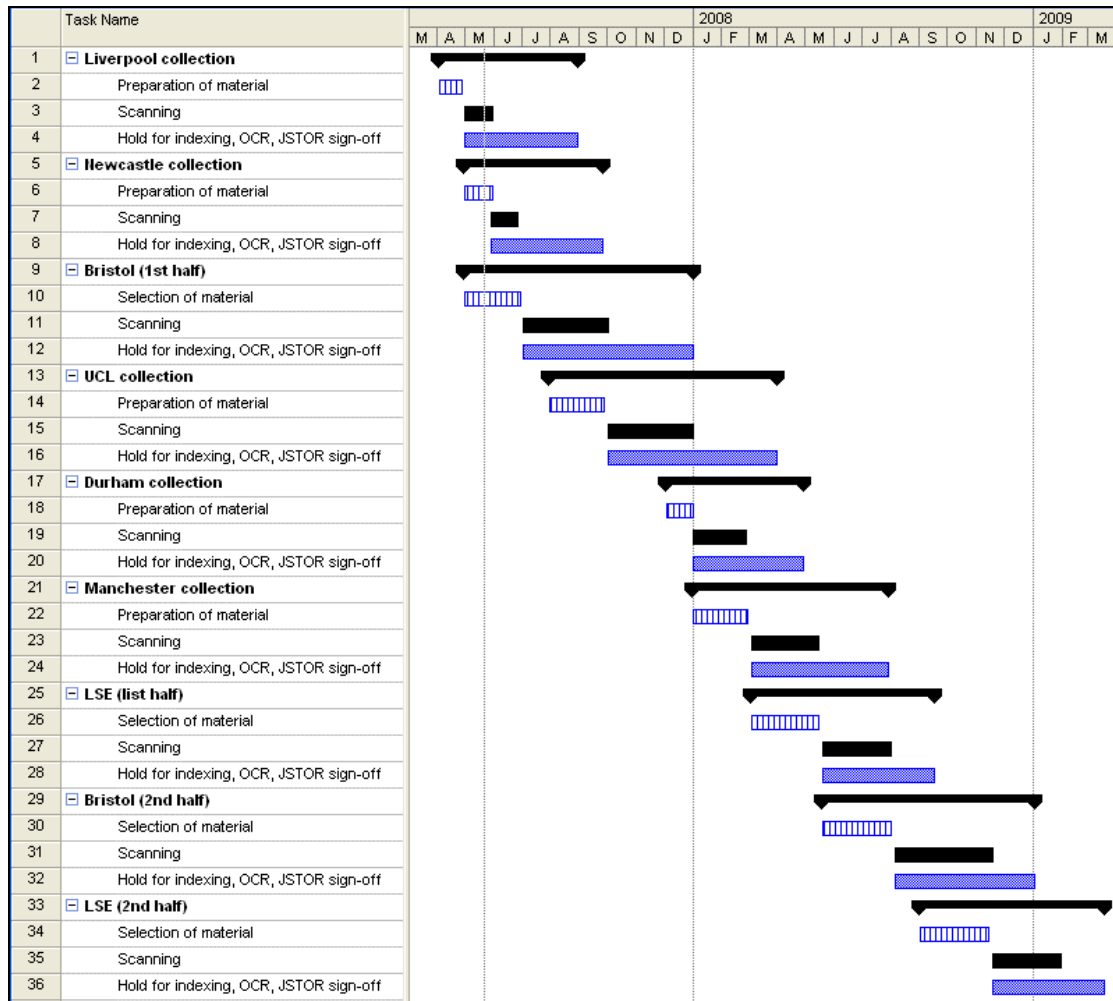
**UK** Repository



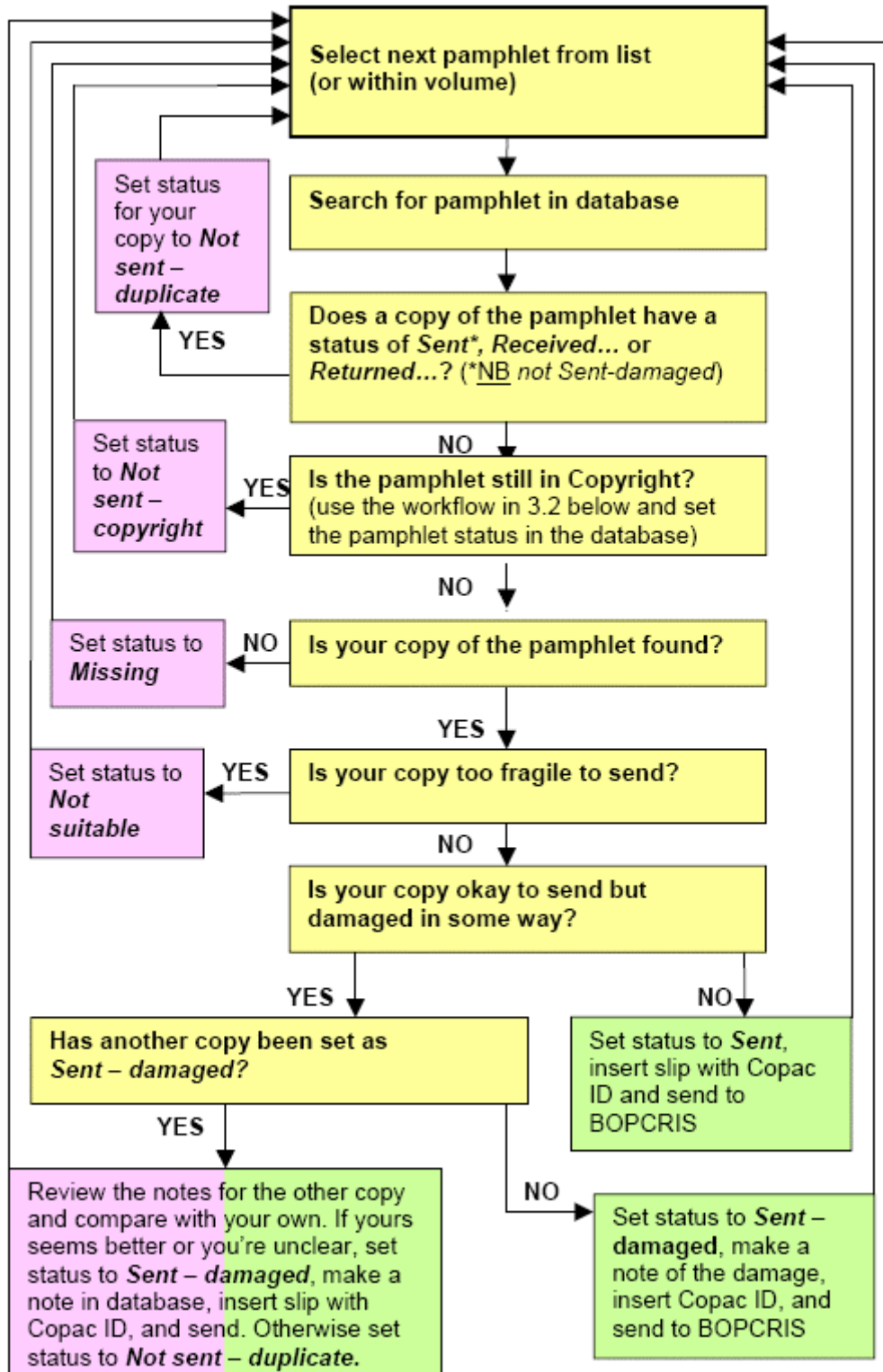
**Gantt chart of Workpackages (WP4-WP10)**



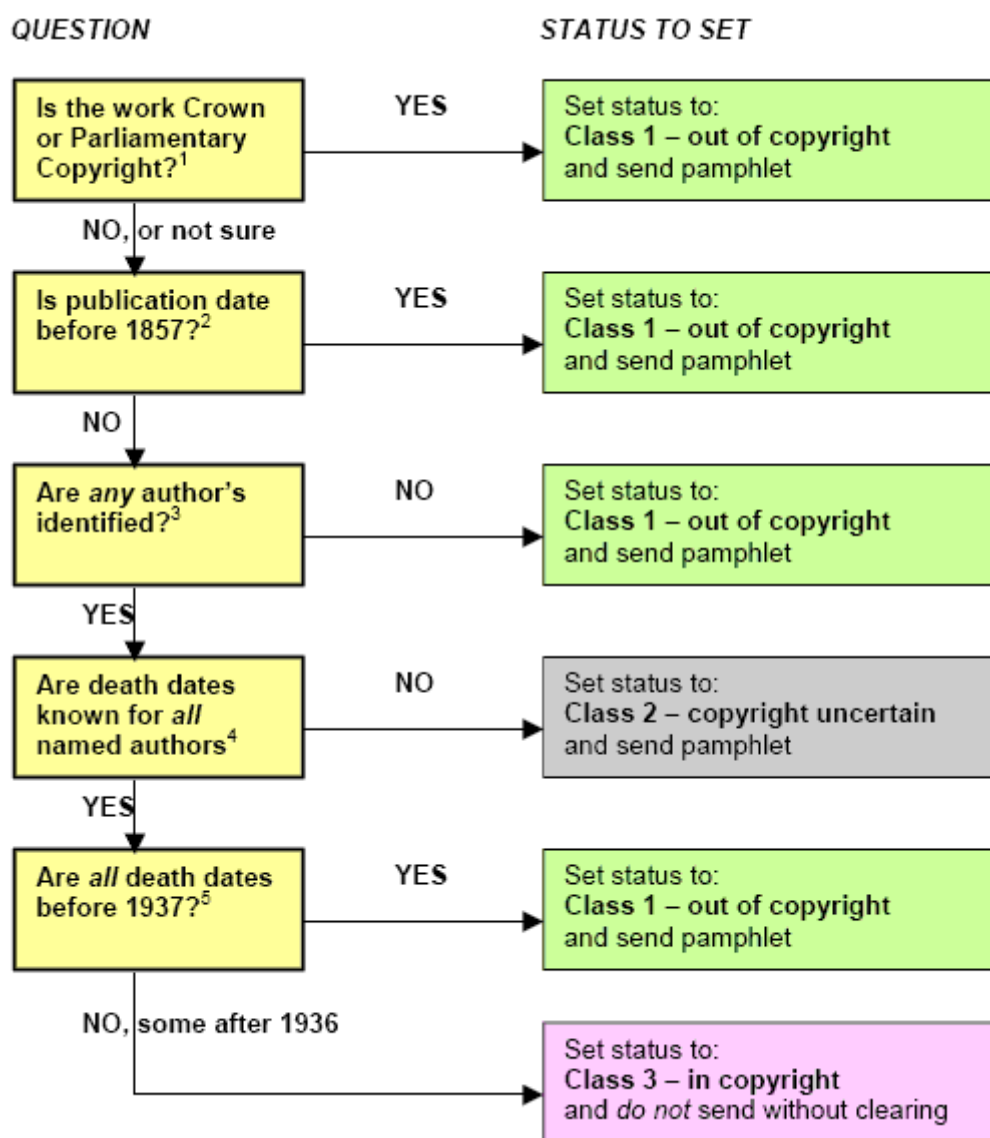
**Gantt chart showing schedule for collections (WP4-WP5)**



## Appendix C. De-selection workflow



## Appendix D. Copyright workflow



### Notes:

1. Crown and parliamentary copyright = 50 years after publication. For definitions of these, see <http://www.opsi.gov.uk/advice/index.htm>
2. This is a conservative cut-off based on the possibility an author may have published at 20yrs and lived to 100yrs.
3. Copyright in anonymous works = 70 years after publication.
4. If an author is named but their death-date is not given, then copyright is uncertain. Libraries may wish to investigate further by (a) checking if any duplicates have dates or (b) checking Library of Congress Authorities <http://authorities.loc.gov/>, Oxford Dictionary of National Biography <http://www.oxforddnb.com/>, or Karlsruhe Virtual Catalogue <http://www.ubka.uni-karlsruhe.de/hylib/en/kvk.html>.
5. If all authors have died before 1937 then the work is out of copyright (e.g. 1936+70yrs=2006). If any author died in 1937 or later then the work is still within copyright and should not be sent without clearance.

## **Appendix E: Licensing agreements**

[Details of licensing agreements have been excluded]