

Project Acronym:
Version:
Contact:
Date:



JISC Final Report

The First World War Poetry Digital Archive

Kate Lindsay, Project manager
University of Oxford

katharine.lindsay@oucs.ox.ac.uk

May 2009

Table of Contents

Acknowledgements	3
Executive Summary.....	3
Background.....	4
Aims and Objectives	4
Methodology.....	5
Digitisation Methodology	6
System Development Methodology	6
Methodology for the Great War Archive Initiative.....	7
User Engagement.....	8
Collecting Requirements	8
User Testing	9
Marketing and Dissemination	10
Implementation.....	12
Overall Project Management and Workpackages	12
The Define Phase – Defining actors and roles	12
The Design Phase – technical architecture and choosing a content management system	13
The Development Phase.....	13
Implementing the Great War Archive.....	15
Outputs and Results	17
Web Site	17
Digital Content	18
Community.....	19
Interoperability	19
Open Source Tools	19
Publicity and Awards	19
Outcomes	19
Realisation of aims and objectives	19
Embedding archival resources in teaching and research – the importance of user engagement.....	20
Allocation of resources – off the shelf v custom build	20
Conclusions	22
Implications	22

Acknowledgements

The First World War Poetry Digital Archive was funded and supported by Phase 2 of the JISC Digitisation Programme. Special thanks are due to the Programme Manager, Paola Marchionni, for her guidance and support throughout the project.

We would also like to thank:

- The project steering group: Duncan Campbell, Richard Emney, Guy Cuthbertson, Paul Groves, Dominic Hibberd, Vivien Noakes, Andrea Peterson, Michael Popham, Simon Robbins, Dunstan Ward, Barry Webb, and Martin Wynne.
- The Literary Estates; The Wilfred Owen Literary Estate, The Robert Graves Trust, The Edward Thomas Literary Estate, The Isaac Rosenberg Literary Estate, The Roland Leighton Literary Estate, The Vera Brittain Literary Estate, The Ivor Gurney Literary Estate, The Edmund Blunden Literary Estate.
- The support and encouragement of all the families of the poets that we digitised.
- The holding institutions who worked with us to identify and digitise material: The Imperial War Museum, The British Library, The English Faculty Library (University of Oxford), The Bodleian Library (University of Oxford), The National Library of Wales, Cardiff University Library, The University at Buffalo, McMaster University, The Berg Collection (New York Public Library), The Britten Peers Foundation, Gloucestershire Archives, The Harry Ransom Research Center (University of Texas).
- The organisations who worked with us to host submission days during the Great War Archive initiative: Oxford City Library, Norwich and Norfolk Millennium Library, Ferens Art Gallery (Hull), National Library of Scotland (Edinburgh), Royal Welch Fusiliers Museum (Caernarfon), and the Powys Archives.
- ProQuest Literature for supplying full text versions of the poems.
- JISC Communications, Flagship Consulting, and the University of Oxford Press Office who helped us raise the profile of the archive in the media
- The encouragement and enthusiasm for the project from Oxford University Computing Services. In particular the NSMS and HFS Services, the Client Relations Team, the TEI, and the Learning Technologies Group.

Executive Summary

The First World War Poetry Digital Archive project has digitised and made freely accessible over 4500 items relating to some of the major British poets of the First World War as well as a wealth of contextual material from the Imperial War Museum and National Archives (audio, video and photographs). In addition the project has launched The Great War Archive, over 6500 items submitted by the general public originating from this historic event.

Intended for researchers, lecturers, teachers, and students across the educational sector, the freely accessible web site brings together items that have previously been hidden away or difficult to access, and has also reassembled dispersed collections. It offers a range of browse and search functionality to facilitate easy item discovery, tools to aid research and teaching, and contextual information to provide a background to the material that it contains. Taking

into account the continuing prominence of the First World War in literature and history courses in compulsory, further, and higher education the archive also offers a range of educational materials to support the teaching and learning of this subject, including seven interactive online tutorials, podcasts, audio tours, resource packs and a portal to other web resources. A Path Creation Tool was also developed and implemented to enable users to create their own annotated trails through the archival contents and beyond.

The processes that the project used to engage the user throughout, and to facilitate the collection of items from the general public, also provide valuable insights for other projects. In addition the web system built to collect and catalogue items submitted by the general public for The Great War Archive has been released as open source for other projects wishing to undertake similar activities.

Background

Between 1996-98, under the JISC Technology Applications Programme (JTAP), Oxford University developed the 'Virtual Seminars for Teaching Literature', based around a multimedia archive holding digital collections relating to the poet Wilfred Owen. Considered pioneering at the time this archive has been one of the major influences on the use of IT in research and teaching within the humanities, and has been regularly cited in publications (e.g. see D. Hibberd *Wilfred Owen: A New Biography*, 2002, the BBC web site). It won several awards (UCISA 1999 'best practice and excellence' in teaching and learning, Oxford's OxTALENT teaching via the web award), and was used by several national bodies (such as the FERL, LTSN, JISC, BECTA, and NetSkills) as a demonstration of excellence. The archive attracted over 1.2 million hits and is referenced by researchers and teachers worldwide.

The widespread popularity of the First World War in literature and history courses across the country is attested to. World War One Poetry and War Poetry in general is taught in most universities, FE colleges, and of course at school level. It also continues to attract considerable public attention due to the cultural importance of the period for modern day Britain, and the way it shaped attitudes to warfare.

In consultative exercises held with users of the existing resource, at workshops, and through direct feedback, it was indicated that expansion of the collection was the primary demand. Therefore this project sought to enhance the resource by broadening the archive to include the digitised manuscripts of other prominent war poets (including Edward Thomas, Isaac Rosenberg, Robert Graves, Vera Brittain, and Roland Leighton), plus relevant contextual materials from the Imperial War Museum (images, audio and video). It was often the case that a poet's manuscripts were spread across a number of institutions, thus by commissioning their digitisation, the project could virtually reassemble dispersed physical collections. Highly aware of the relevance of these materials to teaching and learning across the educational sectors, the project also sought to show how archival resources could innovatively be used in teaching.

Aims and Objectives

The primary aim of the project was to provide online access to a range of digitised material

previously difficult or impossible to access that focused on First World War poetry, and also a range of items to provide context. Material included:

- Poetry manuscripts of Edward Thomas, Isaac Rosenberg, Robert Graves, Vera Brittain, Roland Leighton
- Relevant letters, service records, photographs and other ephemera relating to the above poets
- A selection of film, audio and photographs from the Imperial War Museum
- Items held by the general public relating to the First World War to provide alternative literatures to that of the poets

To achieve this aim, the following objectives were set:

- The migration of all digital content in The Wilfred Owen Multimedia Archive to a new content management system
- The digitisation of approximately an additional 2000 primary source items
- The creation of metadata for each item
- Integration of machine-readable texts of poems.
- The development of a web based interface to search, browse, display and download items.
- The development of a web based interface to collect and catalogue items from the general public.
- Implementation of a path creation tool to allow users to create personalised paths through the archive material. Step by step paths reveal items of interest associated with a particular theme, and notes to prompt further investigation.
- To launch the archive on the 11th November 2008 to mark the 90th anniversary of Armistice.

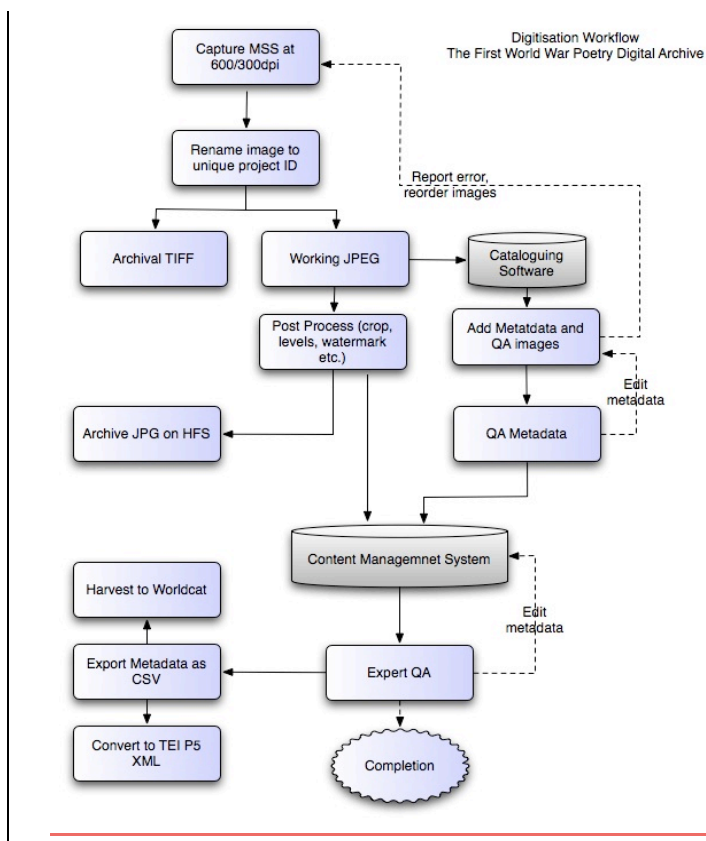
Due to an efficient streamlined digitisation process, reduced costs of digitisation, and the adaption of an off-the-shelf content management system rather than a custom build solution; funds were available to do more than the above. Firstly, additional literary estates were approached and the rights were secured to digitise and catalogue a further 2500 items relating to the poets David Jones, Ivor Gurney and Edmund Blunden. Secondly, the project was able to build a stronger outreach campaign for the Great War Archive initiative, holding submission days at libraries and museums across the country that offered on-the-spot digitisation and advice to the general public¹. Thirdly, the project could concentrate more resource on the development of educational materials to firmly embed the archival resources within a learning and teaching context.

Methodology

The project implemented a series of methodologies to streamline the different processes. A methodology was used for the digitisation of the material and for the overall development and dissemination of the Archive (that implemented user engagement, and marketing). The methodology for the Great War Archive is discussed separately. These are described in greater detail below:

¹ <http://www.oucs.ox.ac.uk/gwa/about>

Digitisation Methodology



Early on the project outlined a process for the digitisation, cataloguing, and migration of all content to be held the archive. Digitisation was outsourced to the holding institutions where source material resided. In most cases these institutions had in place reprographic services that met the project's requirements. In the very few cases where no reprographic services were available (i.e. for personal items held by the poets' families), either the reprographic studio at the Bodleian Library, Oxford University was used or members of the project team digitised the material themselves.

All primary source material to become images were digitised as uncompressed TIFFs, at 600dpi, 24-bit colour, the standard considered appropriate for preservation. From

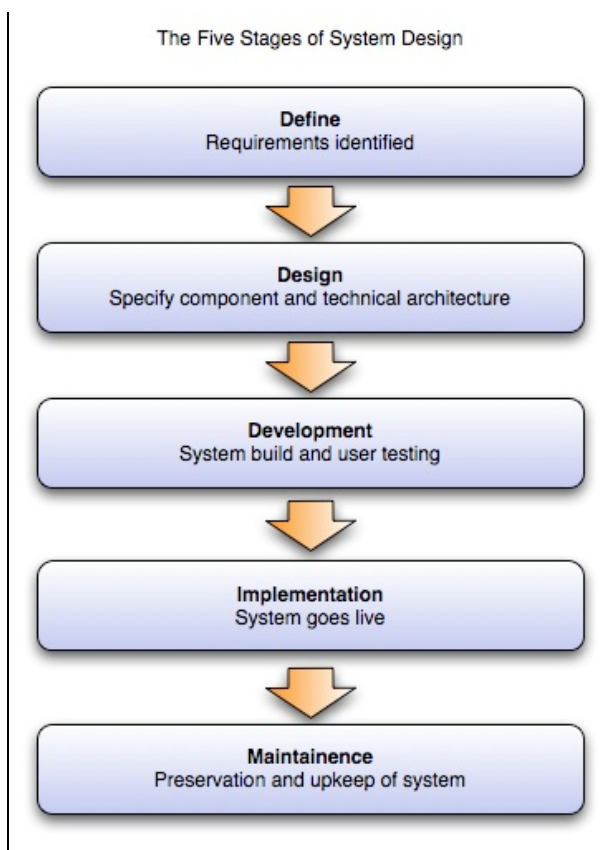
these working copy JPEGs were produced, post-processed and invisibly watermarked using the DigiMarc tools² (so as not to interfere with often detailed content). Post-processing was kept to a minimum to retain the features of the original document. All film footage was digitised as high quality MPEGs and converted to medium and low quality (for web delivery), whilst audio content was digitised as MP3. All high quality content was then archived separately on a Hierarchical File Server for preservation purposes. All images were quality assessed in terms of quality, orientation, etc.

Once content was received and post-processed it was then catalogued using a set of pre-defined fields using the Portfolio software. Metadata for the content was largely descriptive rather than conceptual to avoid imposing subjective meaning upon the items. Once completed this was then quality assessed by the project manager and technical lead and migrated, along with the relevant content, to the content delivery system. Following this each poet collection underwent a process of 'expert QA' by a known scholar in the field of that poet (e.g. the poet's biographer, anthology editor, etc.).

System Development Methodology

The project implemented a five-stage system design methodology to meet the goals of the project.

² <https://www.digimarc.com/mypicturemarc/>



1. The *Define Phase*

During the define phase a set of requirements are identified.

2. The *Design Phase*

During the design phase a range of documentation is produced that specifies the component and technical architecture of the archive. The design is closely mapped to the requirements gathered in the phase 1 and takes account of accessibility and interoperability requirements.

3. The *Development Phase*

Using the documentation produced in the design phase the system development takes place. The design is broken down into individual workpackages that are completed, bug tested and signed off. The development phase also contains periods of user testing, the outcomes of which may mark a return to phases one and two to ensure the needs of the user are met.

4. The *Implementation Phase*

The implementation phase consisted of launch, advertising, marketing.

5. The *Maintenance Phase*

Structures put in place, exit strategy, preservation.

With only a short time period to perform the digitisation and to develop the resource and all its components by the launch milestone, it was felt that a sequential development process would ensure the archive was launched on schedule. However, because of the large and diverse user group the project wished to reach (researchers, teachers and students across the educational sector, and the general public), quite early on it became apparent that continuous user feedback would need to play an important part in the development process. As the project progressed a more agile approach was adopted, with more rapid prototype development. This allowed the project to utilise the feedback received at workshops, steering group meetings, QA sessions, and in formal user testing sessions.

Methodology for the Great War Archive Initiative

The Great War Archive aimed to collect digitised items originating from the First World War held by the general public. The following methodology was put in place for its organisation:

1. Define specific target groups for the Great War Archive
2. Develop an interface for the collection, cataloguing and approval of items via the Web (following the System Development Methodology outlined above)
3. Marketing throughout the initiative period (March - May 2008) focused on the target groups defined in phase 1.

4. Hold a series of ‘submission events’ to perform on-the-spot digitisation for members of the general public, and produce a ‘submission day’ pack for libraries and archives wishing to run their own event
5. Migrate all items and metadata to the delivery content management system

It was decided that material would only be collected via the web and via submission days to decrease workload on the project team (by relying on the public to contribute the digitisation and metadata). By tapping into the potential for digitisation and the public’s interest in the First World War, this methodology also sought to investigate a new model for engaging the public in a University research project, and in doing so the potential for economies of scale of mass amateur digitisation.

User Engagement

The project went through a series of requirements gathering exercises and tests with users prior to launch to check usability and accessibility, and consequently went through a series of refinements. Even now feedback is being collected and usage statistics analysed to allow us to make further modifications. To ensure a usable and popular resource, engagement with users was at the heart of the project. More specific approaches are detailed below:

Collecting Requirements

The first steps to building a set of requirements was to understand how First World War Poetry is currently researched, taught, and studied, the challenges that are faced, and the part that technology and digitised resources can play to enhance this area.

*User-based Steering Group*³

The project formed a steering group consisting of key experts in the field of First World War Poetry. These experts were also researchers and teachers of the subject, and in many cases had worked extensively with the manuscripts that the project sought to digitise. Their experiences in this area were key in deciding exactly what should be digitised, the type of metadata that should be captured, and for generating the requirements of the end delivery system. The steering group approved item lists of digitised material and in some cases visited archives with project staff to select manuscript material.

Teaching WWI Literature Workshop

in November 2007, in collaboration with the English Subject Centre⁴, the project ran a workshop entitled ‘Teaching WWI Literature’⁵. The aim of this workshop was to identify some core themes within the teaching of First World War Literature and in doing so be able to respond to needs, gather requirements, alleviate fears, and illuminate the benefits of using ICT in teaching and learning. The workshop did not focus on technology but instead facilitated discussion on what constituted the effective teaching and study of this area. Discussion questions included:

³ <http://www.oucs.ox.ac.uk/ww1lit/about/steeringgroup.html>

⁴ <http://www.english.heacademy.ac.uk/>

⁵ http://www.english.heacademy.ac.uk/explore/events/event_detail.php?event_index=159

- What are the learning objectives of existing courses you that you teach on First World War Literature?
- Does this differ to what you think they should they be? If so how?
- Should we be teaching courses on War Literature/Poetry or just First World War Literature/Poetry?
- Do the canonical poets truly reflect the attitudes at the time to the War?
- What other “literatures” should we look to include (e.g. drama, film, audio, trench literature, etc.)?
- Are there any practical problems in extending the canon?

Twenty teachers and lecturers attended the event. Discussion sessions were recorded and notes taken throughout the day, providing an important reference point during the define stage of the project.

User Testing

Two extensive periods of user testing were built into the development phase of the project lifecycle and was carried out by the project team and Oxford University Computing Services’ Client Relations Team.⁶ The first stage of user testing took place after the initial development workpackages had been completed to modify the functionality of content management system. This phase looked primarily at the presentation of metadata and the functionality of the content management system itself (e.g. advanced search, image manipulation tools, tools to compare images). Seven researchers were used as testers for this session as this was the group most likely to use the more advanced features of the archive. The second phase of user testing took place after the development of the surrounding web pages and implementation of the browse functionality. Nine testers were used in this phase, a mixture of researchers, lecturers, and school teachers.

The user testing approach was highly qualitative, using a talk-aloud method combined with an online survey⁷. Each participant was guided through a series of scenarios using the archive, and their responses were recorded in the online survey by the evaluators. The latter also took notes on any difficulties the participants had with the functionality, or comments that they had on the content. The results were aggregated and put into order of severity and presented to the project team. Using a matrix that assessed effort, importance and cost requirements were generated and were then either approved or rejected.

Expert QA

Although not originally intended as a form of testing the expert QA process proved very useful as a means of soliciting user feedback. The experts were all key researchers in the field of First World War Poetry, and thus potential users of the archive. The process of using the archive itself to view each manuscript individually meant that we often used some of the archive’s functionality in the process, such as the comparison tool and the zooming tools. Any feedback that the experts gave was noted and fed back to the project team.

⁶ <http://www.oucs.ox.ac.uk/crt>

⁷ Surveys for both periods of user testing and the results of the user testing can be seen on the CRT Wiki at <http://wiki.oucs.ox.ac.uk/ltg-public/WW1projectUserFeedback>

Marketing and Dissemination

To reach the intended diverse user groups of the archive, the project utilised a range of marketing and dissemination methods. These are outlined below:

Web Marketing

The project made full use of web technologies for its marketing and dissemination activities. Within the first month a project website had been set up where visitors could learn about the project's activities and view the latest news and events in the area (delivered via Oxford University's newsfeed system OxITEMS). An 'Image Tasters' blog⁸ was set up and fed into the project website's homepage to give visitors a sample of the types of material to be released on November 11th. Additionally a Facebook group and later a fan page⁹ were set up to disseminate information, as was a Twitter feed¹⁰ which highlighted various items and delved into some of the more hidden aspects of the archive.

To begin to drive traffic to the site, educational materials were released as and when they became available, thus the podcasts (see below) and online tutorials were available for use before the formal launch of the full archive.

Podcasts

Throughout the duration of the project the team interviewed famous commentators and recorded presentations at conferences and workshops (some run by the project itself) and presented these as a series of podcasts¹¹ (recorded via a simple MP3 recorder and edited and compressed in Audacity) In addition, the project also made two educational short videos for download from the website. The two videos were designed to meet the needs of teachers and targeted two important areas – aspects of trench warfare, and tracing your relatives/war memorials. To begin with these were simply fed to the project web site via Oxford's RSS system (OxITEMS) but later these were exposed in iTunes and iTunes U, and the videos also YouTube as part of the JISC channel for the project (this included JISC's showreel).

Friends of the First World War Poetry Digital Archive

To raise awareness of the project, early on links were established with other prominent organisations (e.g. The Western Front Association and The War Poets Association) and individuals (authors and historians). These were named as 'friends of the project'¹², advertising the project on their website, and disseminating materials at their events.

Conferences and Events

The project was represented at a range of conferences spanning both the fields of technology and e-learning (e.g. DRHA 2007 and 2008, Electronic Connections Conference, UCISA) and First World War Literature (e.g. The Annual Robert Graves Society Conference, Siegfried Sassoon Conference, and the Books on the Battlefield Conference). During the Great War Archive initiative talks were also given at numerous Western Front Association meetings to

⁸ <http://ww1image.blogspot.com/>

⁹ <http://www.new.facebook.com/pages/The-First-World-War-Poetry-Digital-Archive/47396039756>

¹⁰ <http://twitter.com/ww1lit>

¹¹ <http://www.oucs.ox.ac.uk/ww1lit/education/podcasts.html>

¹² <http://www.oucs.ox.ac.uk/ww1lit/about/friends.html>

raise awareness. A number of talks were given in-house to raise awareness about the project and promote possible collaborations with other sections of the University.

Public Participation: The Great War Archive

The most intensive period of user engagement was during the ‘Great War Archive’ initiative¹³. The project had the difficult task of reaching a group of users that could, in fact, consist of anybody. It was decided early to focus primarily on three key groups; those who are interested in genealogy and tracing their family history, military memorabilia collectors, and the elderly. Following a user analysis, it was decided to send posters and leaflets for display to all City and County libraries in the country (the mailing list was purchased from facetpublishing.co.uk), along with a letter asking libraries to contact the project team if they were interested in hosting a submission day. Six respondents were chosen and events were held at the following locations:

- Oxford City Library, Oxford
- Norwich and Norfolk Millennium Library,
- Ferens Art Gallery, Hull
- National Library of Scotland, Edinburgh
- Royal Welch Fusiliers Museum, Caernarfon

In addition “submission day packs” were sent to institutions wishing to host their own submission events. This included The Archives of Powys and The Orkney Island Archives. Regular press releases were released during the submission period, both locally in the towns and cities where submission events were being held, and nationally. Family history magazines and military collectors magazines were also targeted.

Later on in the initiative it was noted that Flickr had a number of First World War “groups” where members had already uploaded digitised relics artefacts from the War. By advertising in these groups and contacting members directly the project was able to attract submissions at an international level with material being submitted coming mainland Europe, Australasia, and America.

The Archive Launch

The decision to launch the archive on the 90th anniversary of the Armistice was taken with the view that it would attract more attention towards the resource. However, it soon became clear that this would also mean that the project would have to compete with many other First World War related stories for publicity. The project worked closely with the JISC Communications team and the Oxford Press Office to author and release information to the press. This proved to be a highly successful initiative with publications appearing in over 20 press publications, and numerous online news sites and blogs.

A launch event was held at the Imperial War Museum on the 11th November. Invited guests included prominent scholars in First World War Literature and History, members of the Steering Group, the families of the poets, JISC, and representatives from other related organisations (e.g. The English Subject Centre, The British Library, and The War Poets Association). Invited as an expression of our gratitude for their support over the course of the

¹³ <http://www.oucs.ox.ac.uk/gwa/about>

project, the guests were also gatekeepers of information for potential users of the archive and afterwards would help disseminate our work.

Implementation

The project followed the five stage system development methodology and the digitisation model simultaneously as described previously. During 2008 the Great War Archive initiative was also implemented.

Possibilities for user engagement and feedback were expanded upon at every opportunity, the teachers workshop, user testing sessions and expert QA sessions were built into the project plan from the start and built into the project plan, as was time after feedback had been gathered to revisit requirements and refine the system. Knowledge about the length of time these processes took and the resource required to conduct them was important knowledge when embedding these processes. For instance user testing sessions (to interview c. 8 users and analyse results) took five full days, with a period of 5 full days development work to implement suggestions. Two full days were allowed for expert QA with a following 3 days allowed to modify metadata.

Overall Project Management and Workpackages

To ensure that the project met its milestones and to allow time for important user engagement processes, a series of detailed workpackages were developed using the supplied JISC template. Basic workpackages were implemented from the start of the project and became more detailed throughout the Define and Design Phases. They were revisited at key times, for instance after the user testing sessions, to address user requirements. Two key processes were used to monitor the progression of workpackages. All of the project team were required to submit a weekly report to outline their tasks completed during the week and their tasks for the following week in relation to the workpackages. In the report they were also able to address any risks or concerns they had. A task based report was much more appropriate for this project due to the diversity of the project team which included contractors and graduate students who worked remotely, part time, and full time staff. For the technical workpackages the web-based bug tracking software Mantis¹⁴ was used. This not only allowed for the reporting of issues but also allowed for the input of use cases and milestones and for the project manager to sign off workpackages. Weekly meetings were held between the project manager and technical lead, and bi-weekly meetings with other members of staff. Full team meetings were held every three months to which the contractors were invited.

Contingency was built into the workpackages to allow for unexpected occurrences such as delay in the delivery of images by holding institutions and staff absence.

The Define Phase – Defining actors and roles

The define phase was firstly characterised by the identification of business (project) requirements. These requirements related to the requirements set by the funding body, and also the requirements to migrate the Wilfred Owen Archive/Virtual seminars to a new system. Following this potential users of the archive were defined, and their requirements

¹⁴ <http://www.mantisbt.org/>

gathered. This was done through looking at feedback the old Archive had received, through steering group meetings, and through the teachers workshop.

It was also during this phase that the detailed selection of manuscripts took place and digitisation orders were placed. Greater knowledge of the types of material to be included helped to define requirements. For instance being able to group items together to form a 'compound object' (e.g. a letter or a diary with multiple pages to be presented together in the interface).

From these processes the technical lead developed a series of use cases which defined each major system function in terms of the people or systems participating in the system.

The Design Phase – technical architecture and choosing a content management system

The technical lead evaluated a series of content management systems which could be used to form the backend of the archive. Their functionality was mapped against the requirements documented during the define phase.

The underlying system chosen was CONTENTdm¹⁵. This digital library system allows one to easily link digital objects to metadata for searching and browsing, but also provides a series of tools and utilities for the end-user such as zooming, compound objects, user favorites, and comparison of items. This additional functionality was one of the main reasons why the system was chosen as it allowed the project to deliver functionality that otherwise would have taken significant resource to develop. Moreover the code was open sourced so the developers in the team could enhance the system further and implement important accessibility requirements. Significantly, the metadata could be harvested and also easily output into a variety of formats enabling the future migration to another system if need be.

The decision to use an off-the-shelf content management system negated the need for a full-time web developer post. With the agreement of JISC funds were transferred to pay for the system and its licence and to employ contractors to adapt the system more closely to the projects requirements, develop a user interface, surrounding web pages, the path creation tool, and additional tutorials. This release of funds allowed for much more creative development and design than would otherwise have been able within the remit of the project. Technical specifications and briefs were drawn up to streamline workflow during the development phase.

The Development Phase

Technical Development

System development was carried out in five areas:

1. Conversion of metadata in the old Wilfred Owen Archive to standards and formats used by this project.

The metadata for the items in the old archive had been preserved in SGML format and had to be converted to tab-delimited text for import into the

¹⁵ <http://www.contentdm.com/>

cataloguing software system. This was done via the aid of some simple perl scripts.

2. Adaption of CONTENTdm functionality and interface

Various aspects of the CONTENTdm system needed to be adapted to meet accessibility requirements, tailor search and delivery functionality and also create a seamless look and feel with the surrounding web site. As the code for the system was open to developers this was possible.

3. Development of surrounding web pages

The website was developed using CSS2 style sheets and ruby on rails to embed browse and search functionality.

4. Development of the path creation tool

The path creation tool replicated the functionality of the path tool embedded in the previous Virtual Seminars Project. However, this time the tool was developed in three distinct software components:

- i. The RESTful web service: This is the core of the PCS, handling access to and from the MySQL database and communicating with client software over HTTP. It is intended for use by other software but as it uses HTTP as its transport protocol, can be loaded up in a web browser.
- ii. The generic WUI client: This is a general-purpose, stand-alone web-based client of the web service not tied in any of the specifics of the project terminology or website design. It implements all available functionality of the web service. The generic client was not developed to be deployed on the long-term website hosting but is useful as a sample client implementation for when the service is later made more widely available.
- iii. The project website client: This is another client of the web service, integrated specifically into the project website, and entirely separate to the generic WUI.

5. Integration of the Intute-Lite service¹⁶

The Intute-Lite services enables users to search the Intute database of reviews of Internet resources from within the project website. Search results are returned within the project website's look and feel. Users see a small part of the Intute metadata, and then can click on the resource link or resource title to open the Internet resource itself.

6. Integration of the MyIntute personalisation service¹⁷

The MyIntute personalisation service displays Intute reviews of specific Internet resources that are particularly valuable - organised by keywords which reflect the organisation of the rest of the project's collections and educational resources. The MyIntute service is also used to display links to Internet resources which are not suitable for Intute to catalogue, such as YouTube videos or newspaper articles, and specific pages within a larger resource which are at a too granular level for Intute to catalogue as a separate item.

¹⁶ <http://www.intute.ac.uk/integration/intute-lite>

¹⁷ <http://www.intute.ac.uk/myintute>

Each area was broken down into specific work packages that were signed off on completion. The bug tracking software Mantis was used to track development and the use cases developed were used as reference points for successful development prior to user testing. All development took place on a locally hosted and managed server, using subversion software for version control. When the archive was launched the server was partitioned to provide secure and robust delivery whilst maintaining space to develop further.

Development of Educational Resources

With a direct focus on teaching material and tools the project continued to host the existing online tutorials created in 1998, and 2000, though these were updated to meet accessibility guidelines and migrated across to the new website. As part of this project two further tutorials were created – ‘Computer Assisted Analysis of Texts’ where the user is taken through a variety of online tools that can assist them in analysing the poems (e.g. generating concordances, word clouds, etc). Second, a tutorial was created entitled ‘What is Remembrance?’ that focused on the act of remembrance but drew users into questioning the validity of historical source material.

For teachers who do not have online access in the classroom the team also created a series of resource packs. In short these are zipped files for download containing simple PowerPoint presentations, associated media files from the archive’s collections, and worksheets with further activities.

The project successfully redeveloped the existing ‘Path Creation Scheme’ used in the Virtual Seminars Website. This simple tool was developed as part of the 1996-98 project and allowed users to create annotated paths through the archive for others to follow. This was based on Vannevar Bush’s original concept of trails and trailblazers. The system, however, was very out of date so in this new project it was completely rewritten. Selected paths were then recreated in the new version, and additional ones were added based on the expanded content. The tool is available for users to create their own paths for others to follow as well as following the previously written ones.

It is important to note that for all of the educational materials offered (the online tutorials, the paths, and the resource packs) we made every effort to indicate the level they are aimed at (Key Stage, GCSE, A Level, or Undergraduate). This is not prescriptive but it does give teachers an indication of what they should look at to begin with. Of course all files can be modified to meet the local needs.

Community Development

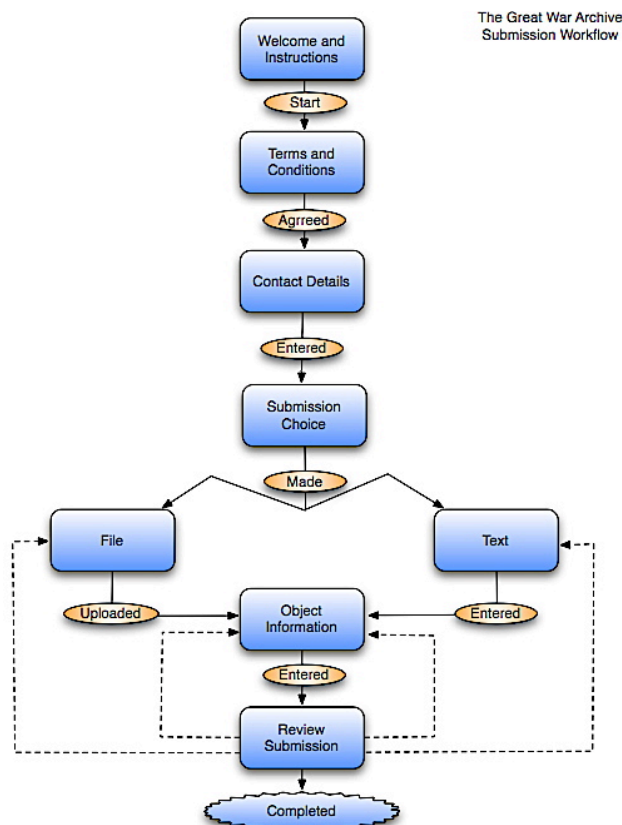
Since 1996 the Virtual Seminars Website had hosted a discussion board for discussing teaching First World War Literature. By 2006 and the onset of the project the software could no longer be supported. The group was moved over to a Google Group, initially moderated by ourselves then volunteers from the group itself. A Google Group entitled ‘Modern War Poetry’ was also set up to capture poetry coming out of modern day conflicts.

Implementing the Great War Archive

The methodology used to design, build and run the Great War Archive initiative also followed the five stages of system design as outlined earlier. However, because the purpose of the Great War Archive was to test the water in regards to finding out what would happen if

the general public were asked to engage in a digitisation project the focus of this aspect of the project was a lot more reactionary and agile.

To build the submission website an initial user analysis was under taken by the project team, much like that performed for the users of the end archive. Types of contributors were brainstormed, along with their characteristics and skills, and requirements derived. The



project also looked at how other organisations had performed similar initiatives such as the BBC ‘Peoples War’ project and ‘Their Past Your Future’¹⁸. It became clear early on that the submission system would need to be as simple as possible to encourage contributions, asking the users to enter minimal metadata but also giving them the opportunity to freely tell the story behind what is was they were submitting. It was also decided not to ask users to create accounts as access restrictions could potentially discourage contribution. The submission website was internally tested throughout the Computing Services Department and feedback gathered via an online survey to ensure a smooth launch on March 3rd. The system was launched within Oxford University for internal submission two weeks prior to its national launch.

Originally scheduled to run between March and May 2008, the high level of public interest and willingness to

contribute, and also the possibility of running additional submission events extended this process until the end of June 2008. At the start of the process the project had anticipated items being submitted only through the web interface and via submission days. However, a large volume of material was posted to the project that needed to be digitised and catalogued by the team with no user contribution. To extend the contribution period and also to digitised posted material posed a significant risk in terms of allowing enough time to QA, post process digital items, migrate to the content delivery system, with significant time left over for user testing. The issue was discussed with the project team, who decided in light of the success of the initiative, to add the extension and digitise the additional material. Resource available from choosing to use CONTENTdm was used to hire graduate students to aid with the metadata QA and digitisation of items, and also to assist with simple technical workpackages, allowing core technical staff to complete the Great War Archive in a shorter time than originally scheduled. Completed in mid-October, the Great War Archive was still able to undergo basic user testing and refinement before the launch in November. Being able to divert resource and allow for extensions was an important aspect of delivering this output.

¹⁸ <http://www.bbc.co.uk/ww2peopleswar/> and <http://www.theirpastyourfuture.org.uk/>

Outputs and Results

The main output of the project was a large web site presenting the user with facilities to search, browse, view, and download over 11,000 items, in a context conducive to teaching and research. However, this simple statement needs further unpacking to fully understand what the project delivered.

Web Site¹⁹

The web site that was launched on the 11th November 2008 acts as the main portal to the collection, providing access to the First World War Poetry Digital Archive and also the Great War Archive.

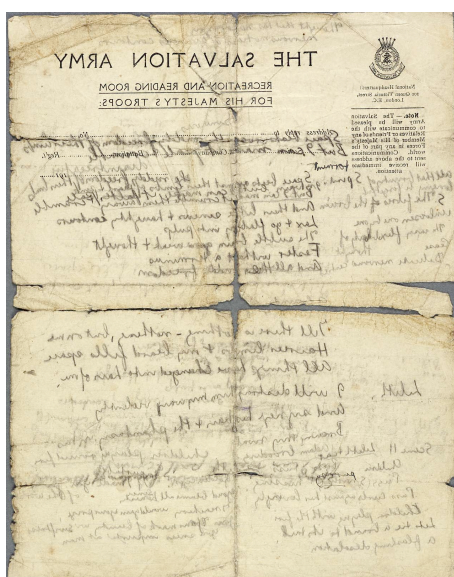
The screenshot shows the homepage of the First World War Poetry Digital Archive. The header features the title 'The First World War Poetry Digital Archive' and a navigation menu on the left with options like 'Search', 'Home', 'The Collections', 'Education', 'The Great War Archive', 'News and Events', 'About', and 'Help'. The main content area includes a 'Welcome' message, a 'Browse the collections' section, an 'Education' section, and a 'The Great War Archive' section. There are also featured links for 'New Oxford Student Text: Wilfred Owen Selected Poems and Letters' and 'Creating First World War Teaching Material'. A 'Featured Path' section highlights 'Highlights of the First World War Poetry Archive' and an 'Archive Showreel' on YouTube. At the bottom, there is a section for 'THE AWARDS 2008' mentioning the Times Higher Educational Awards.

Visitors to the site can choose to search both archives via a series of basic and advanced options, and also browse the material (drill-down) by popular headings such as poet name, title of poem, first line of poem, holding institution, or in the case of the Great War Archive by contributor, content type (e.g. diaries, photographs, letters), or submission day event. Photographs, audio and video from the Imperial War Museum provide a context for the poetry and can be searched via keywords derived from the National Curriculum. The user can build up a favourite list (akin to a shopping cart), browse through a multipage objects item by

¹⁹ www.oucs.ox.ac.uk/ww1lit

item (e.g. page by page in a dairy, or notebook), magnify images, or compare objects side by side (especially useful when looking at manuscript variants of a poem).

All content in the archive is freely accessible for educational use and items can be downloaded and incorporated in teaching materials such as power points and handouts at the click of a button. To support use in teaching and learning the web site also provides an 'Education Area', containing online tutorials, downloadable resource packs, educational podcasts and films, and a gateway to other relevant resources on the World Wide Web that have been catalogued by the Intute service²⁰. The Education Area also offers a 'Path Creation Tool' that enables users to create their own annotated paths through the archival content which can then be published and shared.



Digital Content

The content of the archive consists of over 7000 items of primary source material related to a set of major British poets of the First World War. At the heart of archive lies the poetical manuscripts of these poets. One poem can have many variants and the user can view these to see the workings and processes that the poet engaged with to produce their work. Some of the manuscripts were written on the front line of the battlefields and the muddy, wet and difficult conditions mark themselves clearly on the paper. Along side the poetry manuscripts the archive also contains fascinating letters, diaries, service records, and photographs related to the poets. With the central focus being the war experience of the poet in question, this is supplemented by contextual material (audio, film,

photographs, publications) drawn from the Imperial War Museum and the John Johnson Collection, University of Oxford.

The web site also contains 'The Great War Archive', holding over 6,500 items originating from the War submitted by the general public. Until now many of these items have remained hidden away in people's homes and were at risk of being lost forever. With the launch of the archive they are now freely accessible to the public and help to unlock alternative literatures of the First World War to sit side by side with the Poetry Archive. Submissions included diaries from Gallipoli, love letters from France, audio interviews with veterans, signals sent on Armistice, and a bullet dented tea tin that saved a soldier's life. Every object tells a story, and each story can play an important part in helping today's generation understand what war meant to ordinary people: the soldiers, their families, and the workers back in Britain who kept the country going.

All the material presented on the site is covered by the JISC and HEFCE Model Licence. Even the material submitted by the public as part of the Great War Archive comes under

²⁰ <http://www.intute.ac.uk/>

these conditions, which means that everything can be used worldwide for free by the education sector, though not for commercial purposes.

Community

Through the use of third-party Web 2.0 applications, the project has been able to build a community of users willing to discuss and share knowledge on the First World War. Two Google Groups have been set up; one for discussion on First World War Literature²¹ and one to capture the 'Modern War' poetry coming out of current conflicts such as those in Afghanistan and Iraq²². A Facebook fan page²³ has been created which pulls together all of the Archive's current news and events and also creates a community of 'fans'. Members of the general public can also continue to submit items to the Great War Archive via a Flickr Group²⁴ which also facilitates user-based tagging and discussion.

Interoperability

The metadata held in the system has also been harvested to WorldCat so that objects are exposed in external federated searches. Moreover, as noted in the original bid, all the metadata conforms to the Dublin Core standard and has been also been outputted in TEI P5 XML.

Open Source Tools

The webservice and generic client developed for the Path Creation Tool, and the web system for the submission process for the Great War Archive (called CoCoCo) are all open-source products and will be made available via SourceForge and the web site for others to use and develop under the Open Software Licence V.3.

Publicity and Awards

The First World War Poetry Digital Archive was launched successfully on the 11th November 2008 and was marked by a launch event at the Imperial War Museum, London. An effective media campaign assisted by JISC Communications and Oxford University Press Office ensured that the archive was featured in *The Times Higher*, *Guardian Books* and *The Telegraph* and numerous other publications. The Great War Archive has been highly commended in both the Times Higher Education Awards 2008 and the UCISA 2008 Award for Excellence.

Outcomes

Realisation of aims and objectives

It is noticeable just how much the project has achieved beyond the original scope. The first task was to QA and ingest the Owen material into the new system. Then (as noted earlier) the appropriate items of the named poets and contextual items in the bid were identified and

²¹ <http://groups.google.com/group/ww1lit>

²² <http://groups.google.co.uk/group/modern-war-poetry/>

²³ <http://www.new.facebook.com/profile.php?id=36815907&ref=ts - /pages/The-First-World-War-Poetry-Digital-Archive/47396039756>

²⁴ <http://www.flickr.com/groups/thegreatwararchive>

digitised. However, due to the streamlined processes developed early on in the project, an unpredicted underspend in digitisation, and successful negotiations with the holders of the items (supported by the literary estates and family members) the project could expand its digitisation activities. In addition to the poets identified in the project proposal we were also able to collect similar material for David Jones, Ivor Gurney, and Edmund Blunden (Siegfried Sassoon will also be added as part of the subsequent ‘Enriching the First World War Poetry Archive’ project). The wealth of material delivered is of direct use to research. The collections themselves are dispersed across the world and it would be impossible to place items physically side by side. This project then has led to a major reassembling of these key research resources for the first time since they were produced making this the main resource for researchers working on these poets or this genre. Complimenting these items are thousands of historical items of use to people in different disciplines (e.g. history, military history) many of which had never been seen before (i.e. were held by members of the public in private collections).

Embedding archival resources in teaching and research – the importance of user engagement

It became clear early on that simply making this plethora of content discoverable and viewable is not sufficient when aimed at the educational community. Engagement with users defined certain ways that they wished to both discover and visualise content for educational purposes, for instance browsing via keywords associated with the National Curriculum, hyper-linked metadata, purpose built resource packs to teach particular topics, and being able to present a selection of items with additional notes to students (facilitated by the Path Creation Tool).

We also wanted to make educational resources available via means other than the archive website. For example via lesson planning websites or alternative repositories. As an experiment we trialled uploading resource packs into JORUM²⁵. However, the experience of using the system was not favourable. In 2006 as part of a project funded by the HEA English Subject Centre members of the project team had already looked at JORUM (see S. Lee ‘Depositing a Learning Object in JORUM’²⁶). This proved a useful base to compare with. Although the system had moved on in terms of access (one could now use Shibboleth authentication via the access management federation) there seemed to be little improvement elsewhere. The user interface was still poorly designed and far from intuitive, and the taxonomy used for categorising learning objects heavily biased towards the sciences. For example ‘English Literature’, a major discipline in schools and Universities, was buried under Linguistics. Finally, it was depressing to note that in the intervening three years there was still only two ‘learning objects’ deposited for the entire discipline of English.

Content had to be tied in directly to teaching and learning contexts to be of value, and the project made every effort to do this. User engagement was a vital part of this process.

Allocation of resources – off the shelf v custom built

The decision to use an existing, but adaptable, content management system gave the project much more scope to develop functionality that would otherwise have been beyond its

²⁵ <http://www.jorum.ac.uk/>

²⁶ <http://www.english.heacademy.ac.uk/explore/publications/casestudies/technology/jorum.php>

capacity. In addition to the expected simple and advanced searching functionality, the system also let users create lists of favourites, compare items next to each other, and combine items into 'compound objects' (for instance a poetry manuscript with multiple pages can be viewed in sequence). Furthermore, the CONTENTdm administrative interface allows metadata to be exported into a variety of formats, and also be harvested by the WorldCat Service, thus enabling the project to easily meet JISC's interoperability guidelines.

Digitising and curating collections held by the general public

The Great War Archive also surpassed original estimates. Due to a major national publicity campaign and well-developed workflows and submission systems, it collected together over 6,500 objects submitted by members of the public. Perhaps the most important outcome of this initiative was the investigation of the potential for involving the general public directly in a digitisation project that has great research and teaching value. Lessons learnt can provide valuable reference to other projects wishing to undertake similar activities. The methodology is theoretically reusable in many different contexts. For example, the harvesting of digital items; a mass observation that requires the collection of personal stories; or the submission of teaching materials to build a repository across institutions and disciplines. User analysis to inform system build and marketing campaigns, and different forms of submission mechanisms (e.g. submission events, workshops, web sites etc.) are key concepts that could be reused in many situations.

Additionally what this initiative made clear was the potential for economies of scale that tapping into the potential for mass amateur digitisation could produce. Estimates indicate that each image in the Great War Archive cost in the region of £4.00 to make available, whereas the items in the Poetry Archive cost in the region of £40.00 each (costs include staff time, cost of digitisation, hosting costs etc.). Whilst the quality of the items in the Great War Archive could often be questionable, these figures do support the notion that further investigation is warranted to assess the possibilities of engaging the public directly to build community collections that are of educational and historic value.

Reach and marketing

The use of third party applications or 'Web 2.0 technologies' has been of immense benefit in terms of reaching our target audience. By using social media the project does not just rely on people finding the archive website, it is reaching out to them via their own personal social networks. More and more people are using social media as their gateway to the Web, with time spent on social network websites accounting for 10% of all Internet time.²⁷ Early google analytics show us that the use of social networking tools has not only driven traffic to the web site, but also reduced its bounce rate indicating that we are delivering relevant content to the right groups.

Aside from the technological tools we used to market the website we were fully aware that some groups of users were more likely to find out about the work of the project via other means, such as press publications, reading lists, and information disseminated through libraries and local archives. User profiling early on in the project helped us to identify key dissemination routes and create networks amongst relevant organisations.

²⁷ Nielsen Group (March 2009) *Global Forces and Networked Places: A Nielsen report on Social Networking's New Global Footprint*. Online: <http://tiny.cc/TqyZG>

Conclusions

Even in its early stages the First World War Poetry Digital Archive, together with the Great War Archive, has had a genuine impact upon the teaching, learning, and research of the First World War and its literature. The release of over 11,000 items to enliven lessons, add knowledge and understanding to specific topics, and generate new avenues of research is reaching its target audiences and becoming embedded within academic practice. Both archives are now referenced worldwide in the reading lists and library resources of some of the most prestigious universities, on teaching websites, and by organisations such as the BBC. The Archive's podcasts and educational movies that are available via iTunes and YouTube on completion of the project had reached over 23,000 downloads in five months. Further funding received through the 'Enriching Digitisation' strand will further define the project's impact and continue to reach out to the target audiences of lecturers, teachers, students, and the general public.

This project has also played a part in uncovering the methodologies and processes that are required to embed archival web sites and resources into academic practices. Designing an online archive that is of relevance to many different subject areas and educational levels, as well as different research arenas, poses a number of challenges. By rooting the design of the archive within user engagement activities, and by focusing on a specific set of educational resources that could be used as widely as possible by educators, we have designed a resource that can promote effective use of archival resources within learning, teaching, and research.

Implications

This project has laid the foundations for further work, specifically:

- Public Digitisation. The success of the Great War Archive has shown that involving the general public in the digitisation and curation of collections can produce both collections of worth, and produce economies of scale. The feasibility of this approach needs to be investigated further to assess how the Great War Archive model can be applied to other areas.
- The importance of user engagement. User engagement activities must be built into the project plan from the onset to ensure archival resources are embedded within the practices of their target audiences. To accommodate user requirements development needs to follow a more agile workflow and workpackages need to have contingency built into them. This requires additional planning and resources than would perhaps otherwise be perceived at the start of a project, but user engagement, especially in terms of requirements analysis and user testing, should lead to more focused technical development in the long term.
- Building sustainable communities. Through building established communities of users early on in the project lifecycle the foundations are laid for successful dissemination and also sustainability. Keeping a project 'alive' after funding has ended is always a challenge, but by providing a platform via which users can contribute knowledge and also content can harvest continued interest. However, the monitoring and control of such communities can pose as an issue should institutions wish to control and oversee contributions. The sustainability of communities after funding is an important issue that requires further investigation.

