

## **The Great War Archive: Oxford University's Community Collection**

**Dr Stuart D Lee**

**Reader in E-learning and Digital Libraries**

**University of Oxford**

In this talk I would like to cover some general topics. First, look at a typical digitization project and explore the workflows with this. Then to consider, using the example of Oxford University's project on the First World War, the possibilities of community collections.

At Oxford University for some years now we have been working on digitizing the rare manuscripts we hold in our collections. In the past I was involved in a major study to look at the collections in our libraries, colleges, and museums to consider how we may digitize them and make access possible. One project in particular that I managed was the digitization of the poetry of manuscripts of Wilfred Owen, one of the great British poets of the First World War. This was a typical digitization project, which ran from 1996-1998, and in 2006-2008 we received further funding to work on more poets. In total we photographed 6,000 images, as well as video, and audio, and catalogued these (<http://www.oucs.ox.ac.uk/ww1lit>).

If we consider the workflow process for this it is very typical. We start with the selection of the material, then its capture, archiving, creating derivatives, post-processing, cataloguing, and quality assurance before putting into a search and delivery platform.

To begin with when we consider why we select items for digitization there are many reasons. In short these tend to come down to funding – what someone might pay us to do; but if given a free hand it would include increasing access to the material, strategic or infrastructure reasons, and finally preservation.

However, I would suggest that we need to rethink this whole approach. First of all the relationship between the us – the library, digitizing projects, etc – and them – the user – has changed considerably. If we consider the way users interact with digitized material in the old days they simply would have settled for a searching and browsing front-end web page. However, now the prevalence of Web 2.0 technologies means that users expect far better functionality, they expect the ability to ‘mash up’ or interconnect data, and more importantly they expect to be able to manipulate and contribute to data.

For any digitization project then you have to start engaging with all of these tools and more. Sometimes they may be used for building up communities, or for presenting your material, or for dissemination. Even, one may start to consider using virtual worlds to display your material – such as SecondLife.

Most importantly though is the expectation and willingness of users to create content or contribute it. Let us consider two things. First of all there has been the dramatic change in digitization equipment. High quality overhead digital cameras are still used, there is no doubt, but the most prevalent form of digitization is the consumer camera or mobile phone. This releases considerable potential. If we couple this with what we have observed in the Web 2.0 sphere, e.g. with Wikipedia, it is clear that there is something here we can use. I am not arguing that anything here is new. The concept of communities volunteering en masse to contribute towards public projects is far from new. The most obvious example from my home city is the Oxford English Dictionary.

However, the advances in web technologies and the proliferation of digitizing equipment means that there is now a growing phenomenon that is commonly terms ‘community collection’ or ‘community engagement’. In short, this is where members of the public are asked to contribute material – of any form – to an online collection. This could be images, interviews, or digital stories.

This is where the project I directed comes into play. As part of our wider more traditional project to digitize the poetry manuscripts, we also ran a ‘community

collection' entitled the Great War Archive (<http://www.oucs.ox.ac.uk/ww1lit/gwa>). Here we asked members of the public to send to us any material they personally owned related to the First World War. We advertised the project across the United Kingdom by posters, and other material, and then we invited people to use an online web site to submit their material.

It was important that the submission process should be very simple as here we were engaging with members of the public, not professional librarians. To this end then the metadata or cataloguing was kept to the bare minimum. The workflow also was very simple – a user recorded their name and contact details, agreed to the simple copyright statement, then provided some simple information, and uploaded the files or typed in their text. After that the team reviewed the submission and if it was acceptable we put this into the database.

We also ran submission days around the country where we invited people to bring items to a specific location and we would digitize the material then and there. We developed a system for meeting the people, recording their information, and scanning it in.

The material we received was quite extraordinary. More importantly this included the stories associated with the objects, the personal recollections behind the medal, letter, diary, etc.

During the three months we received over 6,500 items from the public – items which would not have been seen otherwise. This engaged the public in the project, but more importantly was very cost effective. In the traditional process we found that each item digitized cost around £40.00, whereas in the community collection it cost less than £3.50. This was because if we consider the workflow the cost on the project had been greatly reduced. We did not select the material, or digitize it when it came to the online submissions, or catalogue it.

On the whole we found that the quality of the material was good, the cataloguing was not a problem, and that this was a model that could be followed elsewhere.

At Oxford we have now been funded to assist others to create community collections. Under our Running a Community Collection project (RunCoCo - <http://blogs.oucs.ox.ac.uk/runcoco/>) we are training libraries on how to run a similar project, as well as making available the software we used free of charge as open source. This will allow anyone to set up a collection site, define the metadata to be requested, check each submission, and then offer a simple browsing and searching engine. In effect a cradle to grave solution.

2014 will also mark the hundredth anniversary of the start of the First World War. As part of this we are hoping to be part of a major EU project led by the Europeana consortium that will run a community collection across Europe training various countries to run their own Great War Archive. However this is just the beginning.