

# CENTRE FOR HISTORY AND ECONOMICS

## The Digitization of History Meeting with steering group at the Computer Laboratory August 11, 2008

We were delighted to welcome steering group members David Palfrey, Michael Roe, John van Whye, Melissa Lane, and Steven Murdoch. In a wide-ranging conversation, we discussed some of the following points:

JV thought that we should make our mission statement clearer (perhaps changing the name of the project to reflect this) and that we should feature a site about other similar projects in the UK (AHDS and JISC) in order to clarify our 'place' among other digitization initiatives. ER pointed out that we should continue to highlight our interdisciplinary nature and the group's focus on the digital experience 'outside North America'. We are also planning on undertaking some small-scale digitization projects of key resources at Cambridge of interest to historians.

This led to a discussion about some of the challenges facing archives outside of the US. LD reviewed the group's previous discussions on the inability of OCR to handle Arabic and Indic-language scripts (largely due to a lack of development of alphabetic keys). ER spoke about the problems surrounding basic conservation issues in India. Mismatches between digital indices and physical collections however, if not addressed systematically by historians, may lead to serious biases in future research. RW remembered Natalie Ceeney's suggestion in her Easter term talk that perhaps 70-80% of the archive remains largely unresearched. There is potential for digitization to allow a reassessment of kinds of bibliographic search tools which historians have followed into the archive up to now, but we may lose the chance at reappraisal if we wait too long.

ML here noted that we also need to address the unwillingness of historians to get more involved in digitization (low attendance at previous talks; Natalie Ceeney's point that TNA had been unable to involve historians in consultations over their new system). We may need to redefine the project slightly to draw in larger audiences and get more historians interested in the issues. JV wondered whether we should focus some talks on 'best practices' for historians or on 'how to make a great discovery' using new research methods. ML suggested that we make the links section of the website more 'active' by linking to projects featuring in talks and group meetings, in particular featuring TNA's discussion wiki. She also suggested that we continue and perhaps expand on the plan to podcast either audio or video clips on the blog and website. ER noted that we could start this off with a feature on copyright encapsulating some of the points that MR raised about the differences between UK and US copyright. LD wondered whether some of the reticence could be put down to ignorance of the issues and the possible consequences for freedom of access for historians in the future; perhaps some of SM's points about open publication at Harvard should also be a focus for the group (i.e. not just Digitization of History but Historians and Digitization?) One of the group's mandates might therefore be to provide historians with tools to understand the transformation of research, i.e. tips on 'how to do history now'. ER felt that the possibilities for new methods of access to the French notarial archives and Indian archives are both prime examples of how '21<sup>st</sup> century history' is quite different.

SM called attention to Geoff Goodell's article on the internet presenting a different 'face' depending on location of access, a point which further reinforces concerns that transferring print copyright standards to the digital world may simply worsen existing inequalities of access. SM suggested that we get in touch with Wendy Seltzer at American University. MR noted his experience with Project Gutenberg's copyright clearances mimicked those employed by Google in terms of greater restrictions outside the US, and ER mentioned work on IP regimes and 'open

knowledge'. The steering group seemed unanimously in favour of a workshop to address some of the implications of copyright clearances, and of a talk and podcast by Mike Roe.

We began and concluded with discussion about the blog, the consensus generally being that it should provide brief overviews of interesting projects and speakers, or short video/audio 'podcasts' that reference more detailed articles on the website or elsewhere, rather than attempt to replace the CHE static website. The key thing is for both the website and the blog to see frequent updates (no less than every 1-2 weeks for the blog) and to try to keep things as current as possible with talks throughout the terms.

LD & RW, August 2008