ncse nineteenth-century serials edition

Newsletter

Issue 1: June 2005

Welcome to our first Newsletter. This will be a bi-annual posting that will provide details of the objectives, progress, and achievements of the project.

Funded by the Arts and Humanities Research Council, **ncse** is a collaboration between Birkbeck College, King's College London, and the British Library. Its remit is to develop an exemplary edition of six nineteenth-century newspapers and journals in electronic form. The titles are:

- *Monthly Repository of Theology and General Literature* (1806-1838).

- *Northern Star* (1838-1852).

- *Leader* (1850-1859).

- *English Woman's Journal* (1858-1864).

- *Tomahawk* (1867-70).

- *Publishers' Circular* (1880-1890)

ncse will feature full-text facsimiles, including adverts, illustrations, and other supplementary material where possible. Metadata and concept maps will allow sophisticated searching and indexing both within titles and across the edition, opening up this material in new ways that the user can manipulate according to his or her needs. Challenging conventional uses of print media by emphasizing its interconnectedness, the project provides both a new technical and methodological model for electronic editions and a new conceptual framework for researchers.

"The Newspaper Editor," *Illustrated London News*, 11 September 1880.

The Project Team

ncse is a collaboration that draws on the expertise of its team members and the resources of their institutions. Not only is the British Library offering access to its unparalleled resources but it is also providing us with newly-produced microfilm for scanning. The Centre for Computing in the Humanities (CCH) at King's College London is undertaking the technical implementation of the project. Its experience in overseeing large-scale digitization projects, as well as its keen awareness of the intellectual issues involved in creation of these resources, complement the specialist knowledges offered by the research team.

Director: Laurel Brake, Birkbeck College.
Team: Isobel Armstrong, Birkbeck College.
               Marilyn Deegan, King's College London, CCH.
               Ed King, British Library.
               Jim Mussell, Birkbeck College.
               Suzanne Paylor, Birkbeck College.
               Harold Short, King's College London, CCH.
               Simon Tanner, King's College London, CCH.
               Mark Turner, King's College London.

Early Stages

The project officially began work on 1 January 2005. After introducing ourselves to each other, the first weeks were spent establishing institutional relationships. As a result of this early activity, the **ncse** office in the Faculty of Continuing Education at 26 Russell Square was equipped with furniture and equipment, additional workspaces were provided at the British Library's sites at Colindale and St Pancras, and access to these resources was arranged for the project's research assistants.

Even though the project is only publishing six titles, this still represents a large amount of material. To avoid becoming overwhelmed by print – and as exasperated as the newspaper editor overleaf – the team spent the first two months familiarizing themselves with the journals. Brief surveys, which we called "snapshots," were taken of each title in order to register any major changes of format, personnel, content etc. This information was captured in a table which evolved as the team began to encounter more complex configurations of data. These snapshots thus provided the templates that underpinned the next exercise: page turning. Page-turning was envisaged as a long-term process, designed to further familiarize the research assistants with the journals, the state of the holdings, and any problems these might present the optical character recognition process (OCR). As work commenced, the team quickly found that this method suited those publications that appeared in a regular format, such as the *English Woman's Journal*, but was substantially more laborious for titles that did not maintain a consistent structure across numbers, such as the *Northern Star*. As a result, the team redefined the goals of page-turning: as well as using it as a record of historical changes in the journals, we also used the process to validate the outcomes of our next task – the production of data-maps.

Data-mapping marked the project's first interaction with the technological demands of our software. The production of structural analyses, which would map the relationships between classes and levels of data, raised significant methodological issues for the research team. Periodicals, as a print genre, are predicated on change over time, with readers being offered novelty within clear structural constraints. Identifying these, while still allowing the flexibility to capture both the differences between titles and the changes within titles, was a demanding task. For example, leading articles, which are recognizable features across many (but not all) of the titles, are often in different positions and have markedly different purposes. Although the team felt leading articles were a consistent unit, we were wary about over-determining both their role and their importance. Structural features such as divisions and departments also provoked discussion. Should departments be assigned descriptive labels that might apply across all the titles? How should the major division between front and back, which delineates the types of content in weeklies such as the *Leader* (and so changes the meaning of those departments that migrate across the division over the run), and is mimicked in the satirical *Tomahawk*, be captured in our structure?

As we are approaching the time when our digital material will become available, the research team have begun to delimit the categories of data that will allow searching across the edition. We have three levels of data: the first is the automatically-generated demarcations that we have encoded as part of the OCR process; the second, which we have called advanced metadata, consists of index terms such as people, places, publications, institutions and societies, and events; and the third, which provides the

thematic terms for our concept maps. Whereas the data-mapping exercise informed the first level of data, levels two and three are being developed by the team through a series of meetings. At these, sample pages are interrogated according to candidate terms in order to both gauge their efficacy across different types of content and also estimate the time needed to mark-up accordingly. For the advanced metadata categories we will use a combination of manual mark-up and automated processes. Marking up according to the concepts will be more complex: not only will the bulk of them have to be identified by a human operator, but the relationships between concepts must be carefully evolved from the periodicals themselves. These meetings then serve an interesting dual process: they allow us to address the practical considerations determining our encounters with the material while also granting this material agency in shaping the meta-level structures that will determine users' interactions with the edition in the future.

Over the next few months the iterative exchanges between the research team, CCH and Olive will increase as the digital material begins to appear. Further intellectual work will have to be done as software is tailored to our needs, and our interpretations of the archive are reflexively conceived according to the demands of the technology. What began as a complex print genre, but with recognizable units in articles, numbers, volumes, and runs, has rapidly become much more elaborate. The rich dialogic exchange between our historical knowledge and the possibilities offered by new technologies has forced us to reconsider, often at quite fundamental levels, many of our assumptions about this material, and the uses to which it can be put.

Conferences and Seminars

In order to raise the profile of the project, team members are participating in a wide range of conferences and seminars. We are eager to publicize the interim findings of our research, as well as learn from our peers and explore further the domains in which the project might find users. As well as participating in major events in nineteenth-century studies and periodical studies, the team have also been attending and participating in events in history (including media history and history of science), computing in the humanities, and material culture.

The team recently consolidated its links with the Centre for Nineteenth-Century Studies at Birkbeck College. The Centre brings together scholars both from across the College and also from other universities with which it has relationships. This interdisciplinary ethos is shared by **ncse,** and Suzanne Paylor and Jim Mussell presented a paper at the Centre's internal symposium in May. This paper, "Mapping the 'Mighty Maze:' the Nineteenth-Century Serials Edition," detailed the pressing need for digital archiving of paper-based resources, while also considering the intellectual demands of the periodical press and the ways in which this archive is reconfigured by the application of new technologies. This paper will be published in the first number of the Centre's online journal, *19: Interdisciplinary Studies in the Long Nineteenth Century*.

In July the research assistants are travelling to the first NINES (Networked Interface for Nineteenth-Century Scholarship) workshop, hosted at the University of Virginia's Institute for Advanced Technology in the Humanities. The expenses of sending two scholars to the United States have been met by a combination of funding from the University of London's Convocation Trust and NINES's own bursary scheme. The event promises to be stimulating: not only have the organizers brought together representatives from a number of projects in the United States and Britain, but they also offer a series of events in which we can benefit from the resources and experience of NINES.

Later in the year members of the project team will be at events focussing on book history ("Material Cultures and the Creation of Knowledge," University of Edinburgh) and the history of science (British Society for the History of Science Annual Conference, University of Leeds, and the History of Science Society Annual Meeting, Minneapolis). Media history – obviously a key component of our activities – has also been well represented: Mark Turner and Laurel Brake attended Media History and History in the Media at the University of Wales, and will be at the Research Society for Victorian Periodicals Conference in

Washington. Members of the team also participated in the first meeting of the Leverhulme International Media History Network. In September Jim Mussell will present the project to the second Mapping the Magazine Conference at the Cardiff School of Journalism, Media and Cultural Studies. In fact September is our busiest month in terms of presentations. In one week the two interests of the project coincide: ncse will present a panel "Nineteenth-century periodicals in the long view" at the British Association of Victorian Studies, which will be chaired by Isobel Armstrong and feature papers by Mark Turner, Laurel Brake, and Jim Mussell; meanwhile Suzanne Paylor will present the project to the annual Digital Resources in the Humanities conference in Lancaster.

<u>Aims and Objectives</u>

The team believe that participation in as many fields as possible is vital to sustain the project both in its development and after its completion. As collaborative research projects become more common in the humanities, and the funding bodies increasingly direct their attention to electronic resources, it becomes essential that we recognize the diversity of potential users, and also the related research undertaken in fields that might be parallel to our own. The team have attended seminars at CCH, introducing us to a wide-range of projects that have encountered similar issues to us despite radically different source material. We have also begun to engage with institutions such as the London Knowledge Lab, a collaboration between social scientists and computer scientists from the School of Computer Science and Information Systems at Birkbeck, and the Schools of Mathematics, Science and Technology, and Culture, Language and Communication, at the Institute of Education. By publishing the edition online, we move beyond the academy: much of our outreach work is aimed at learning from others who already work in such contexts, and to identify specific contexts of our own.

It is through these encounters that we aim to embed **ncse** within the work of our peers and colleagues: not only will the resource radically alter users' encounters with the periodical archive, but we also hope that it will provide conceptual and methodological tools that they might find useful. In reconfiguring our six titles for electronic publication we also change them, facilitating new ways of encountering the rich material that the superfluity of print might otherwise occlude. Our deep-level mark up exploits both the automated processes of our optical character recognition (OCR) software as well as the specialist knowledge of the researchers on the team. The combination of search facilities that we are developing – concept maps, advanced metadata, and full text keyword searches - will allow users to navigate this structural foundation in sophisticated and entirely novel ways. It is in navigating this plenitude of material that the project is original: in a landscape of mass digitization projects, we offer the means for users to refigure the cluster in ways that are meaningful to them.

Jim Mussell

Nineteenth Century Serials Edition
FCE
Birkbeck College
26 Russell Square
London
WC1B 5DQ
Tel: +44 (0) 20 7631 6693
Fax: +44 (0) 20 7631 6270

We are keen to hear from those interested in the project. If you have any material that we might include in the next **ncse** newsletter then do contact us.

Visit our website at www.ncse.kcl.ac.uk.