Nineteenth-century Serials Edition Project (NCSE)

NCSE is a three year AHRB funded project which, in collaboration with the British Library, seeks to achieve two key objectives: first to establish a unique TEI compliant electronic edition of six nineteenth-century periodicals and make this freely available, via the web, with historical and critical apparatus and extensive research functions, and second to provide a conceptual and technical exemplar to be utilised by future projects.

Rationale:

Researchers across the humanities conducting studies which seek to utilise nineteenth-century periodicals currently face manifold difficulties that prohibit the fullest excavation of these rich and important source materials. Currently, the resource base suffers from three primary difficulties: firstly the periodicals are fragile in their paper form and in many cases are already deteriorating; secondly the resources are geographically dispersed across research libraries with limited access; and thirdly the resources are currently static, in that they are unsearchable and benefit from no contextual/historical facility.

The novel and innovative approaches made possible by newly available OCR technology have the power to resolve these problems and throw open a range of resources for a new generation of research across disciplines. The outcome of NSCE will be first to provide what will be a valuable resource in its own right -- a unique electronic archive of periodicals with metadata and 'concept maps' for users from multiple disciplines, emphasizing the theoretical, cultural and historical dimensions of the edition. Second NCSE will function as an exemplar to future projects in developing appropriate innovative and practicable methodologies and procedures for such work.

Challenges:

The successful fulfilment of NCSE's key objective, to create not only an electronic resource of lasting value but also a model for future projects, requires that the project team consult with experts in the field and potential resource users at every stage of development. The project team feel strongly that NINES will offer us a unique opportunity not only to begin this consultation process in earnest, but also to form lasting cooperative networks for the creation and exchange of information and expertise with specialists familiar with the peculiarities and specific challenges posed by engaging with nineteenth-century media using twenty-first-century technology.

The 2005 summer workshop coincides with what will be a particularly pivotal point in the progress and development of NCSE. At this time we will be starting to confront in an informed manner the practical and conceptual challenges posed by our project. Having made an initial assessment of the source materials, and digitised approximately 4000 of our 30,000 page total we will have a good impression of the practical problems involved in creating our OCR text. Furthermore, we will have begun to form the concept maps, which, through software implementation, will provide the structural and semantic features of our data. In July 2005 NCSE will be optimally poised to both contribute to and benefit from the excellent forum that the NINES workshop will provide.

Plans and progress:

This is a three year project, funded by the AHRB and in collaboration with the British Library that commenced on 3 January 2005. The project team, Professor Laurel Brake (Birkbeck, Project Director), Professor Isobel Armstrong, (Birkbeck), Dr Mark Turner (King's College London), Dr Suzanne Paylor and Dr James Mussell, (Researchers), Prof Harold Short, Simon Tanner, and Dr

Marilyn Deegan (all CCH, KCL), and Mr Ed King, (British Library), are currently engaged in the assessment of source materials to identify any specific problems which may have an impact on the project timetable or modus operandi. After three months we will begin to research the structured metadata for the archive. In the second half of year 1, we will begin to identify the conceptual and thematic terms which will act as organizing mechanisms for the structured metadata. These will be reviewed with the International Advisory Group and others at the end of year 1. In year 2, with nearly a third of the source material digitalized, we will have a working 'concept map' of the structured metadata, which will then be further developed by team members. By the end of year 2, we expect the bulk of the archive to have been digitalized and approximately two-thirds of the structured metadata to be determined. The technical team are currently engaged in the scanning and processing of the source materials with Olive Software's Active Paper Archive (APA) software. The teams will then work to specify and develop the XML DTDs required for the scholarly support materials; to design the overall architecture of the project web site, taking particular account of the technical demands in integrating Olive's APA software with the support materials; to specify and develop the XSLT required to transform the support materials for display on the web, including generation of indices and search forms; and to specify and develop the procedures needed to integrate the various technical components. The technical framework will be based on the DTDs of the Text Encoding Initiative (TEI), and will draw on wide experience in a number of CCH projects with these DTDs. CCH also has significant experience with Olive's APA, through its involvement in similar projects. In year 2 the key technical work will be to enhance and refine the tools developed in year 1, and to develop them to ensure smooth integration of the journal materials in Olive's APA with the scholarly support materials. In year 3 the key technical work will be to refine and enhance the technical structures and processes in the light of the feedback from the RAs and Project Directors, and from members of the Advisory Group and other scholars invited to submit comments. An important element of year 3 work will be the collaborative development, involving the whole project team, of a full description of the development and delivery model on which the system is based, covering the associated editorial issues, technical approaches used, standards adopted and the tools and procedures developed..

Support:
Harold Short, Director, Centre for Computing in the Humanities at KCL, has the role of Technical Director in the project, overseeing the technical development work and reporting to the Project Director. The technical development work in the project will be fully supported by the range of expertise available in CCH, which includes extensive experience with and expertise in text mark-up and analysis, web development for scholarly purposes, and relevant database and image expertise. The AHRB through the Arts and Humanities Data Service Executive (AHDS) and its Service Providers, provides guidance and regulations for all the projects in which it is involved. The guidance and advice of the OTA (Oxford Text Archive) on the project is ensured by inclusion of an OTA representative on the Advisory Group. Furthermore, advice has already been sought, and will continue to be sought, from the British Library, the Text Encoding Initiative Consortium, the International Advisory Board and members of related projects. Participation in the NINES summer 2005 workshop would greatly enrich the project's development.

15 Jan 2005

Birkbeck, University of London
26 Russell Square
London WC1B 5DQ
UK