Manuscript Information	
Journal Acronym	**RJVC**
Volume and issue	
Author name	
Manuscript No. (if applicable)	**_A_461286**

QUERIES: to be answered by AUTHOR

Dear Author
Please address all the numbered queries on this page which are clearly identified on the proof for your convenience.
Thank you for your cooperation

QUERY NO.	QUERY DETAILS
1	Please explain WebCT in full on first use (only use).
2	Please explain acronym in full on first use.

DIGITAL FORUM

Reading in the Digital Archive

Shafquat Towheed

In the century that formed the mass reading public, the figure of the reader was a ubiquitous and all-pervasive presence. Reading alone, reading aloud, the reader as a public orator and spectacle, the female reader, reading while travelling by train, reading in the public library, the reader as a subject of intense surveillance: visual images and textual representations of the act of reading were conspicuously prevalent in the arts, culture and commerce of the nineteenth-century. Both quantitative and qualitative scholarly studies have provided us with compelling evidence of the massive increase in the range, frequency and variety of reading that took place in the period.[1] Not only were readers consuming an unprecedented volume and range of reading material, but they were recording their widespread engagement with textual matter, often (as in the case of upwardly mobile autodidacts), for the first time. The nineteenth-century material archive is rich with the evidence of reading practices that readers have left behind: commonplace books, marginalia, accounts of reading in correspondence, memoirs, autobiographies and diaries, the aggregated data of lending libraries, working men's institutions and reading groups, the disaggregated data of inquisitorial records, commissions, ephemera and advertising, and a wealth of visual representations (pictorial, illustrative and photographic) of reading. The cumulative richness of this material, filtered and interrogated synchronically and diachronically, provides valuable information about reading in the nineteenth-century, but how do we digitize the nineteenth-century material archive for our own digital century, while maintaining valuable (often implicit and extra-textual) information about historical reading practices as they happened?

---

1. See for example, William St Clair, *The Reading Nation in the Romantic Period* (Cambridge: Cambridge University Press, 2004); Stephen Colclough and David Vincent, 'Reading', in *The Cambridge History of the Book in Britain, Vol. 6, 1830–1914*, ed. by David McKitterick (Cambridge: Cambridge University Press, 2009), pp. 281–323; Martyn Lyons, *Readers and Society in Nineteenth-Century France: Workers, Women, Peasants* (Basingstoke: Palgrave, 2001); and Simon Eliot, 'From Few and Expensive to Many and Cheap: The British Book Market, 1800–1890', in *A Companion to the History of the Book*, ed. by Simon Eliot and Jonathan Rose (Oxford: Blackwell, 2007), pp. 291–302.

Books stripped and scanned

Mass digitization projects have been systematically scanning nineteenth-century
archives, often before a concerted policy has been codified. The *Internet Archive's*
<http://www.archive.org> ever-increasing digitization of out of copyright nine-
teenth-century material (821,786 items as of 1 November 2009) offers the largest
selection of scanned books from the period, as well as potentially a wealth of
information (in the form of marginalia, book plates, marks of ownership, and
borrowing stamps) about how these books might have been read, during the
nineteenth-century and after. However, this ancillary data is disaggregated and
accidental; it has not been tagged, recorded or marked-up, and cannot be searched or
examined in a systematic way (there is for example, no way of finding marginal
comments in a scanned volume, or looking for specific book plates or evidence of
ownership). The critical mass of digitized material already achieved by the Internet
Archive and by Google (*Google Books* <http://books.google.com/>) presents some
very particular opportunities and challenges for scholars working on the nineteenth-
century. The free availability of scanned, out of print and difficult to find primary
printed sources means that academics can be increasingly adventurous in terms of
their pedagogy, setting for example, works by authors that have since fallen out of
favour, but which are often more representative of the period than well-known
canonical writing. This has implications for classroom study, the use of the set reader
or anthology, and the specific types of student engagement with the text itself. In the
near future, nineteenth-century literature, history and culture courses administered
largely through the medium of a virtual learning environment (utilizing *Moodle*
① <http://moodle.org/>, *Blackboard*, or other WebCT technologies) will begin to
depend heavily upon this extensive back catalogue of digitized material, almost
certainly at the expense of printed anthologies or readers. Tertiary level education in
② the humanities increasingly foregrounds competence in ICT as a key skill set, and so
we run the risk of teaching comprehensive searching and evaluative skills, without
offering students sufficient awareness of the difference between their own reading
practice and that of the historical period being studied. This is especially important in
the case of the rise of mass literacy (and with it, a mass reading public) in Britain; as
teachers, we neglect the patent differences in reading practices between a society
acquiring literacy, and one moving to digital literacy, at our peril. Nor is such a
transformation without its technical challenges. If nineteenth-century readers
struggled with inadequate lighting, broken type, and small print, twenty-first-
century readers must contend with the vagaries of optical character recognition, the
bottleneck of data download restrictions, and the unwieldiness of graphically
intensive web-pages. Indeed, with regard to the clarity of optical character
recognition (an essential requirement for the successful scanning of Chinese
characters), the cost of scanning, and the production of new (mainly mobile) digital
infrastructures, we may, as with the birth of movable type, have to look to East Asia
for direction.

   As the digitization of texts proceeds apace, the staples of nineteenth-century
reading practice – the circulating library, the periodical press, and reading aloud

within the family circle – become ever more difficult for us to explain to contemporary readers, for the material evidence is increasingly degraded. Most disturbingly, the progress of *Internet Archive* and *Google Books* and the guaranteed online access to pre-twentieth century works that they promise provides an economic rationale for the deletion of considerable extant library holdings. Public libraries are already in the process of selling or destroying large swathes of their nineteenth-century fiction and newspaper holdings; with the removal of this material archive, we lose valuable additional evidence of nineteenth-century reading practice, as well as significantly reduce the familiarity of contemporary readers with the size, shape, format, smell and texture of these books.[2] For the new generation of e-readers, now more familiar with the digital scroll than the printed codex, the very materiality of the extant, carbon-based nineteenth-century archive will present considerable challenges of utility, access, and comprehension. Utilizing digital resources such as *Internet Archive* will offer students and scholars of the future a wealth of easily accessible texts, but it will not by itself equip them to undertake productive research in the material archive. Thus while expertise in the skills set for e-reading will increase exponentially, we may end up dematerializing the very texts that form the foundation for a solid intellectual grounding in nineteenth-century studies, and taking students further away from the intrinsic material aspects of reading practice in the nineteenth-century. This has potentially dire consequences for training graduate students in bibliography and the history of book, for example. The problem of the free availability of content (for most of human history, the major limiting factor in the educational curriculum) is being increasingly replaced by the widespread anxiety over the generation of context. How do we know which book to read, how to read it, and what kind of context in which to place the book, its author, and its readership? Nineteenth-century readers would have benefited from considerable paratextual and extratextual material in shaping their choices of what to read, and how to read it. Advertisements and lengthy publishers' lists in the end papers of volumes would have given them a sense of how each book sat within a wider context, while the vibrant reviewing culture would have given many readers a clear sense of the appropriate selection and sequence for reading a set of interrelated books. The doctrinaire categorization and classification of reading material in the nineteenth-century libraries, as well as the sometimes systematic policing of who had access to what, predicated particular reading practices; and of course, the prevalence of the multi-volume book (especially the three-volume novel) dictated a specific sequence for reading. None of these sets of valuable additional information are reflected in the mass textual digitization programmes such as those carried out by Google or the Internet Archive; there is no way at the moment, for example, of systematically

---

2. See for example, John M. Morrison, 'Lovely New Library, Shame About the Books', *The Guardian*, 28 December 2006, <http://www.guardian.co.uk/books/booksblog/2006/dec/28/lovelynewlibraryshameabout> and Sarah Cosgrove, 'Waltham Forest: Council Admits it Destroyed Books', *The Guardian* 29 November 2007, <http://www.guardian-series.co.uk/news/1866925.0/>

locating the advertisements or publishers' lists bound into the end papers of so many volumes of nineteenth-century books.

Journals and Newspapers

With a particularly high density of tagged metadata, Gale Cengage's *19th Century British Library Newspapers* (also known as *British Newspapers 1800–1900*) < http://newspapers.bl.uk/blcs/ > provides a wealth of accurate searchable functions, opening up remarkable possibilities for humanities researchers. Some two million pages of content can now by systematically searched in seconds, the results can be tagged, and the identified files saved, printed or e-mailed. However, it is evident from the design of the interface that the database is primarily designed as a highly accurate and responsive retrieval tool, and not as an electronic repository; it does not encourage or reward browsing (the term is not used once in the descriptive information introducing the resource), in stark contrast to the *actual* use of the hard copies of these newspapers in the homes and reading rooms of the nineteenth-century. The difficulty of translating broadsheet pages to the entirely different aspect ratio of the computer screen means that the full page as presented in condensed form is often illegible, and the default presentation is designed to focus on the specific articles of columns that had been explicitly sought out through the detailed search. This means full, double page scanning and browsing (the staple of nineteenth-century newspaper reader) is very difficult to achieve. The prospect of browsing the pages of your chosen newspaper is further compromised by an automatic log-out for inactivity, which is linked to your mastery of the sequential (one tagged article after another) or searchable reading favoured by the digital resource. Paradoxically, while the digitization of nineteenth-century newspapers means that they have never been more widely available, the very experience of casually browsing through the pages of an essentially disposable publication has become increasingly remote and difficult to reconstruct.

Similarly, while nineteenth-century newspaper readerships would have been carefully delineated by cost, geographical location and the size of their print run, modern mass digitization projects present a level playing field of newspapers, with each competing title receiving the same level of attention and occupying the same amount of visual space (in terms of data storage and their implied position within the structured content of the database). Ironically, the imperative to be comprehensive in projects like *19th Century British Library Newspapers* undermines a deeper understanding of the realities of newspaper reading in the period, which was neither truly democratic nor egalitarian in terms of its access to reading matter, but rather typified by the limitations of distribution, the exigencies of tariffs and taxes, the relentless push and pull between local and national news, and a fierce level of competition between titles for readers. The plethora of newspapers amassed together and digitized in totality in *19th Century British Library Newspapers* belies the fact that no individual or group of individuals living in the nineteenth-century would have even been able to scan the list of titles available in the database, let alone read them comparatively, something we expect to be able to do through the advanced search

facility. In presenting undifferentiated data without a hierarchy or additional ancillary information, mass digitization projects partially re-contextualize the reading matter of the nineteenth-century. They do not offer us an accurate reconstruction of how newspapers may have been read, but instead present readily searchable decontextualized content.

Back to the Future? The Codex vs the Scroll

Finally, while digitization proceeds apace and a new generation of online, potentially mobile e-readers are beginning to explore the books of the past as well as their own times, the issue of the personal intimacy of the reader with the text becomes increasingly acute. While readers have for centuries experienced the physical intimacy engendered by an immersive reading of printed books – we continue to talk about 'curling up with a book' – the equivalent idiom for the on-screen reading of digitized material has unsurprisingly, yet to be coined. Although it remains to be seen if future readers will curl up with Amazon's Kindle, I do not want to say that such a relationship will not develop; the astonishing attachment of users of 3G mobile phones to their handsets suggests that an intimate relationship with texts delivered through digital means will emerge. Whether these texts will be presented in the form of a pseudo-codex (as with Kindle), or a new digital version of that old form, the scroll (as with reading tools for mobile phones) remains to be seen, but either way, the distance separating the twenty-first century digital archive from the nineteenth-century material one will inexorably continue to grow.

Shafquat Towheed
Open University
S.S.Towheed@open.ac.uk