

## **A Word with our Critics: Doing Media History with Electronic Resources**

The title of this paper is not a self-defensive rhetorical gesture on my part, but rather is the title of an occasional note published in the mid-nineteenth century weekly *Tomahawk*. *Tomahawk* was a satirical newspaper, published between 1867 and 1870 that rivalled its more-familiar contemporary *Punch* for both wit, political commentary and high-quality cartoons and caricatures. “A Word with Our Critics,” which appeared in October 1867, seemingly defends *Tomahawk* against critical opinion. “It is no use disguising the fact;” it reads, “we have few friends in the literary world. We are reviled and spat upon for telling the truth; we are called contemptible for not truckling to the conventionalities which are the yoke under which others are content to grind.” (Anonymous 1867: 229). This robust defence of their editorial style seems straightforward – it must be a response to unfavourable reviews elsewhere in the press – but in a self-reflexive journal like *Tomahawk*, such assumptions are premature. Not content with mocking nineteenth-century society, *Tomahawk* also spoofed the periodical as genre, including fake reviews (complete with fake concerts, novels and plays), leading articles, correspondence, and even advertisements. As such, “A Word with Our Critics” might not refer to any publication at all, rather it could be a satire on similar articles in the wider press.

Such ambiguity presents us with familiar historiographical and bibliographical challenges. To determine whether the article is a response to a piece in another publication, we must begin to read a vast number of possible titles, often in poor condition, probably scattered across a large number of different museums and libraries.

Online public catalogues can help identify extant runs, but the relevant information remains buried in an abundance of titles, pages, articles, and words. There are a number of resources that help navigate the archive, and some of these – particularly the *Waterloo Directory* – are available in easily searchable electronic forms. However, like many of its contemporaries, *Tomahawk* presents significant textual challenges: it was not signed, so authorship is not a useful delimiting category; it contains a variety of different content, in a variety of genres, requiring a breadth of knowledge; and its textual form invites complex reading strategies, both for its readers today and those in the past. What *Tomahawk* demonstrates is that context – both in terms of its own run, wider nineteenth-century resources, and the various interventions that have been made since – is vital for meanings, but that such contexts can also overwhelm the researcher in an abundance of print. “A Word with Our Critics” exposes a further dimension of textual uncertainty, inscribing ambiguity into historical analysis.

*Tomahawk* summarizes some of the wider issues that I will address in this paper. It is unmistakably embedded within its historical context: not only does its form embody its modes of production, but its advertisements, commentary and reviews reveal its place both within the print trade and the wider contexts of nineteenth century culture. As a Saturday weekly, *Tomahawk* – and indeed other well-known nineteenth-century titles such as the *Athenaeum*, *Punch*, *Illustrated London News* and the *Saturday Review* – is positioned on the border between the magazine and the newspaper, commenting on passing events, but also publishing more reflective essays and reviews for weekend reading. As such, *Tomahawk* comments on events as they occur rather than the more

stabilized moments of history, and its satirical gloss renders already unfamiliar events and figures even more obscure through its tantalizingly oblique allusions. Its visual material, which invokes its own satirical repertoire, requires that we complement lexical analysis with a way of reading that can delimit the relationship between word and image on the page.

I suggest that it is the interrelated aspects of ephemerality, material form (including periodicity), historical specificity, and cultural interconnectedness, that present the most robust challenges to our efforts to produce digital resources that both conserve and make accessible our surviving print archives of magazines and newspapers. An important part of the work of the Nineteenth Century Serials Edition (**ncse** for short) is the interrogation of the periodical as a genre. This afternoon I would like to outline some of the challenges involved in republishing historical archives of magazines and newspapers as digital editions. I will begin by discussing the generic features of serial publication, and argue that any functional digital archive must recognize the complexities of material form as well as providing accurate textual transcripts. Despite the ephemeral nature of the periodical press, it represents a vital archive for research into historical culture. Yet its conservation also requires scholarly intervention in order to ensure that the appropriate care is applied to the digitization of these idiosyncratic print forms. Digital archives, I suggest, only go so far: it is essential that we recognize the specific demands of this material, and devise the appropriate scholarly and technological solutions to open them up in ways that do not overly determine them as texts.

\* \* \*

The project, now eight months into its three year span, will produce a digital edition of six very different nineteenth-century periodical titles, complete (where possible) with advertisements, prefaces, and supplements. In order to navigate this diverse material, we are embedding metadata to allow searching within a range of delimiting fields, and also implementing concept mapping, to facilitate research within configurations of conceptual and thematic categories. The six journals are taken from across the nineteenth century, and represent the diversity of forms and periodicities within the genre. Alongside *Tomahawk* is the *Monthly Repository*: a metropolitan magazine, at first Unitarian but increasingly non-sectarian, that contains literary commentary, biblical criticism, and significant early nineteenth-century poetry; the *Northern Star*: a heady mix of mid-century working-class radical political commentary, international, national and local news, poetry, and entertaining tit-bits which was also the organ of the Chartist movement; the *Leader*, a progressive, liberal and secular weekly established by Thornton Lee Hunt and George Henry Lewes; the *English Woman's Journal*: a mid-century, middle class monthly that embodied the feminist goals of the Langham Place group; and finally the *Publishers' Circular*: a long-running, London-centred, fortnightly trade publication for the print and publishing industry.

Taken as a cluster, these journals challenge our conception of the periodical itself. Whilst the *English Woman's Journal*, an unillustrated monthly containing essay reviews, poetry, correspondence, and commentary, fits comfortably into our notions of the nineteenth-

century periodical, when considered with its advertisements its political goals become explicit, and the text functions as a portable adjunct to the employment agencies situated at Langham Place. We have deliberately included two newspapers within the cluster, exposing the often arbitrary nature of the division between newspaper and magazine. The *Northern Star*, for instance, is self-consciously a weekly newspaper, but it also published poetry, moving this towards the beginning of its letterpress as the run progressed; the *Leader*, whilst also considering itself a weekly newspaper, has an even more ambiguous identity, containing two discrete sections, one concerning news and current events, and a second, positioned at the back, containing literary commentary and essay reviews. By publishing these titles together, we both gesture towards the discontinuities inherent within the genre, while also foregrounding its centrality for the nineteenth-century print trade.

We view the published result of the project as an edition: we are republishing these six journals for a readership today. This raises significant methodological challenges: unlike contemporary electronic publications, we must maintain fidelity with the paper-based sources as the accuracy of how we render them in electronic form determines their scholarly value. This means that the edition that we are creating cannot exploit digital technology in the same way as other online versions of print media. Although we are making the edition freely accessible online, and so are capitalizing on the internet's ability to provide multiple versions of our edition simultaneously across the world, we cannot participate in the same temporal dynamic as other online publications. **ncse** is also an archive, and so in some ways we reproduce book publication as we will publish a

single static edition in one go. However, it is not enough for us to simply reproduce the words on the page, we must also reproduce the periodicals' spatial arrangements, pagination, typography, images and other graphic devices. These are all items of scholarly interest, and so must be identified and marked as objects that are available for manipulation by users. Periodicals, of course, are not books, and so the edition must also be structured in a way that reflects the diversity of reading strategies that are inscribed into the structure of article, department, number, volume and run. Just as we are aware that our own scholarly biases might tend to over-privilege certain forms of content – literary or political for instance over sporting news or puzzles – so too must we avoid over-privileging certain structural units. Nineteenth-century readers did not read whole pages from cover to cover in a number, and neither will our users. Our edition of six titles thus becomes rapidly more complex, as its finite number of pages becomes fragmented into a vast number of structural units, all available to be reconfigured according to the demands of the user.

In order to manage this large amount of information, we are currently implementing four levels of data recovery to complement simple free text searching. These four levels – segmentation, structural metadata, advanced metadata, and concept mapping – address the limitations of browsing the facsimiles or searching the transcripts. The processes through which the text is generated from the microfilm source is not 100% accurate, and the size of the archive limits the extent to which we can correct it. By using fuzzy searching the software, to some extent, compensates for the inevitably inaccurate transcription. However, for some of our titles, this is not sufficient. The *Northern Star*

for example, being a cheap and rather flimsy working-man's weekly, is in poor condition for the application of the technology that produces the textual transcript. It was poorly printed, and smudges of the impression, tears, and creases often obscure the letterpress. In addition, the poor quality paper on which it was printed has deteriorated markedly, and so the edges of pages and quality of the letterpress have deteriorated further.

It is the privileging of text over form, however, that represents the most significant limit of free-text searching, no matter how sophisticated. Unable to identify images, some typographic fonts, or spatial layout, such searches do not directly reflect what is on the page – in fact they explicitly ignore it – and so take little account of the contextual references that structure meaning. The first level of markup, segmentation, directly addresses this problem. The research team, in collaboration with Olive software and our partners at the Centre for Computing in the Humanities at Kings College London, are developing ways in which we can use imaging technology to recognize and demarcate formal features on the page. Preliminary trials with Olive have suggested that it is possible to recognize and map items on the page in ways that maintain the complex relationships between them. For instance, a title on a volume wrapper often differs from a title on the first page of an individual number, necessitating the recurrence of the metadata item title at a number of levels. This diversity within the periodical runs results in complex hierarchical nests, and we are adding structural metadata to the segmented items in order to grant them independence as objects while recording their structural location.

The third and fourth levels of markup exploit the xml that is produced by the imaging process in order to provide more interpretative means of navigation. At the third level of markup, advanced metadata categories map indexical terms throughout the edition, allowing users to trace their occurrence across and within the titles. We have identified seven main categories: people, places, events, institutions, products, publications and businesses. Although we recognize the overlap between these categories – a publication can be a product, and a person can be a business – they are designed to delimit data in order that it can further be explored. We are currently working to construct a database, built from extant lists of data, to contain this information in a way that we can manipulate to overcome some of the limits of free-text searching. For instance, much of the nineteenth-century press is pseudonymous: by marking up all people, it becomes possible to link authors to their pseudonyms, perhaps uncovering the more surreptitious activities of authors and editors in their texts. Equally, by identifying place names, we can situate them within a hierarchy – for instance identifying the “Strand” as a place important for the print trade, but also as a district within a larger entity called “London.” These categories also allow us to engage with the indirect allusions within titles like *Tomahawk* and also non-textual content such as the lavish images in the *Publishers’ Circular*

The fourth – and we’ve imagined it as the highest – level is concept mapping. This is an interpretive search tool which will operate by identifying thematic content. In a series of discussions we have identified conceptual terms and linked them in a three layer hierarchy. These concept maps will allow users to reconfigure the edition according to their research interests: for instance those interested in the urban will be able to delimit

all items that deal with the experience of the urban by selecting the master category environment, and the subset urban. Again, the scope of the edition limits the extent to which we can interpret content and mark up accordingly: although we are exploring key word extraction, lexical databases, and semantic linking, computers cannot read, and so are no substitute for expert interpretation. We are still in the process of developing both the conceptual and electronic tools to implement concept mapping, but the objective is to provide flexible search methodologies that will bind the edition together in unorthodox ways. As the concept map is both evolved from the journals, but not dependent upon them, we hope that it will be both a valuable research tool in its own right, and also – as a product of a specific moment in nineteenth century studies – become a historical artefact. By mapping the edition thematically, we hope to produce a new perspective on this material, complementing historical narratives with an emphasis on the interconnectedness of nineteenth-century print culture.

These four levels, and the technical architecture that structures them, thus respond to the challenges of material form, historical specificity, cultural interconnectedness, and ephemerality that I outlined earlier. By providing digital images of the pages as well the facility to perform complex searches on the generated text, the edition will render these titles more accessible than either their nineteenth-century paper equivalents or a twenty-first century electronic transcript would alone. Simultaneous searching across the edition – by formal feature, metadata item, or concept – will emphasize its nature as cluster and, while we make no special case for the interdependence of these six titles, demonstrate the inevitable cross-references inscribed within it. Indeed, such serendipitous encounters are

central to research into the periodical press: the volume of material insists that research will never be exhaustive and, just as before digitization, there will still be an element of chance to what users find. What **ncse** does is to map data fields in a way that opens them up, offering users the chance to reconfigure the edition while conserving the historical integrity of the paper-based archive.

It is only through recognizing the malleability of the periodical archive, and then incorporating the digital tools to preserve this in its electronic form, that we can remain true to the bibliographic challenges of the press. Periodicals are more open texts than books, employing a wider visual repertoire and inviting a number of reading strategies in their efforts to reach readers. It is thus important to sustain these features in any electronic version, even if this radically increases its complexity. **ncse** recognizes the contextual nature of the press – that it does not exist in isolation at the level of title, department, article, or even word – and inscribes this within its internal structure. Not only do digital technologies offer us an opportunity to explore historical artefacts in new ways, but there is an increasingly pressing need to engage with their politics. Increasing numbers of periodicals are being offered electronically: Proquest, Thomson-Gale, and the British Library are all currently engaged in large-scale digitization projects, but their scale prevents scholarly attention to the product, and their commercial goals limit the access to the resulting resources. **ncse** is a research project that not only identifies and delimits the generic features of the periodical, but is also developing the electronic tools to offer all internet users unprecedented access to all its contents.

**Works Cited**

Anonymous (1867) "A Word with Our Critics.," *Tomahawk*, 1: 229

Dr James 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

<http://www.ncse.kcl.ac.uk>