THE DIGITAL TURN

Bob Nicholson

To cite this article: Bob Nicholson (2013) THE DIGITAL TURN, Media History, 19:1, 59-73, DOI: 10.1080/13688804.2012.752963

To link to this article: http://dx.doi.org/10.1080/13688804.2012.752963

Published online: 04 Jan 2013.

Submit your article to this journal

Article views: 1335

View related articles

Citing articles: 2 View citing articles
THE DIGITAL TURN
Exploring the methodological possibilities of digital newspaper archives

Bob Nicholson

Advances in digital technology have made the recent past seem like a foreign country. Media historians did things very differently in 2002. In the last decade, hundreds of historical newspapers and periodicals have been digitised and made available to researchers via online archives. Whilst the emergence of these resources has generated contrasting responses from historians, an increasing number of researchers are now embracing the new methodological possibilities created by keyword-searchable digital archives. As the first examples of this scholarship begin to appear on the horizon, this paper considers whether media history is on the cusp of a ‘digital turn’. It outlines the existing responses to digital methodologies, deconstructs digital newspapers in order to explore how they differ from their paper originals and uses case studies drawn from my own research into the late-Victorian transatlantic press to demonstrate how new methodologies might be applied.

KEYWORDS digitisation; newspapers; digital methodologies; transatlantic; American slang

Introduction
The rapid digitisation of newspapers and periodicals has transformed even the recent past into a foreign country. Historians did things differently in 2002. At that point, a relatively small number of titles were available online. The Internet Library of Early Journals offered free access to limited runs of six publications, ProQuest’s Periodicals Contents Index had migrated from CD Rom to the Web and Gale was about to launch the pioneering Times Digital Archive. Reflecting upon the emergence of these resources, the Victorianist Patrick Leary predicted that the digitisation of press archives would become one of the most ‘exciting and far-reaching developments in . . . the twenty first century’ (Leary ‘Digital Age’ 204–5). It is remarkable how quickly this vision became a reality. Less than a decade later, thousands of English-language newspapers, magazines and periodicals are available online. Major titles such as the Illustrated London News, the Scotsman, the Financial Times and The Economist are all available in individual databases. However, these projects have been dwarfed by the emergence of multi-title databases such as Gale’s 19th Century British Library Newspapers (71 titles), 19th Century UK Periodicals (180), 19th Century U.S. Newspapers (500) and 17th and 18th Century Burney Collection (1271); ProQuest’s, Historical Newspapers (36) and British Periodicals (460); and non-commercial projects such as Chronicling America (686), Nineteenth-Century Serials Edition (6), New Zealand’s Papers Past (70) and Australia’s Trove (255). This extraordinary growth shows little sign of slowing down. At the time of writing, the recently launched British Newspaper Archive (BNA) holds more than 180 titles and is expanding at a rate of 8000 pages per day. When
combined with other online archives such as Project Gutenberg, Eighteenth-Century Collections Online (ECCO), the John Johnson Collection, the Database of Mid-Victorian Illustration, Google Books and the Internet Archive, this adds up to a significant amount of print culture—all available online, all searchable by keyword. Historians have never had such a wealth of information at their fingertips.

It is important to stress that these resources have not been distributed equally. Copyright concerns have limited the digitisation of twentieth-century newspapers and periodicals. Major papers such as The Times, the Illustrated London News, the Daily Mirror and the Daily Express are all available post-1900, as are most of the titles included in ProQuest’s Historical Newspapers database. However, these publications add up to a mere fraction of the total available to historians of earlier periods. Whilst the BNA now provides access to a small number of provincial newspapers which span the late-nineteenth and early-twentieth centuries, the majority of multi-title databases end their coverage in 1900. Digital archives of seventeenth- and eighteenth-century print culture, on the other hand, have been marred by problems of quality rather than quantity. As James E. Tierney has argued, the inconsistency of the period’s typography reduces the accuracy of scanning software, whilst an overreliance on materials accumulated by contemporary collectors has left large gaps in the archive (Tierney). The British Library’s Burney Collection, for example, draws upon newspapers and pamphlets collected by the Reverend Charles Burney (1757–1817); whilst it appears to contain an impressive 1271 titles, most of these are not available in complete runs. Key archives of eighteenth- and twentieth-century print culture are also characterised by poor interfaces; UK Press Online, which holds the Daily Mirror and Daily Express, is, for example, a near-unusable mess which fails to implement even basic features such as hit term highlighting.

Nineteenth-century newspaper archives fare better in almost every respect. The absence of copyright concerns has allowed a larger number of provincial and metropolitan titles to be digitised. Moreover, these texts have been sourced from systematically compiled archives rather than private collections. As a result, the presence of missing issues has been significantly reduced (though not eliminated entirely). The improved quality of nineteenth-century typography also results in more accurate keyword searches. A study commissioned by the British Library found that the average ‘word accuracy’ of their nineteenth-century newspaper database was 78%, compared to the Burney Collection’s 65% (Tanner et al). This is not to suggest that nineteenth-century newspaper archives are perfect, rather that they are qualitatively and quantitatively superior to those available for other periods. Whilst the methodological debates outlined in this article might fruitfully be applied to archives of eighteenth- and twentieth-century print culture, the following discussion is focused primarily on nineteenth-century databases. It is here—where our tools are at their most powerful and our materials are at their most abundant—that debates on the methodological potential of digitisation are most likely to bear fruit. The first section evaluates how Victorianists have responded to the methodological implications of digitisation and discusses the prospect of a ‘digital turn’ in humanities scholarship. Part two deconstructs the digitisation process and unpacks the concept of a ‘digital methodology’. Finally, the concluding sections draw upon my own research on transatlantic media relations to demonstrate how both qualitative and quantitative digital methodologies can be applied to the field of media history.
The Digital Turn

Historians have responded to the emergence of online archives with cautious enthusiasm. Much debate has surrounded the pitfalls and drawbacks of digital research. Leary has warned of an overreliance on online archives and the dangers of an ‘offline penumbra’; an ‘increasingly remote and unvisited shadowland into which even quite important texts fall if they cannot yet be explored [or identified] by . . . electronic means’ (Leary ‘Googling’ 82). Some researchers have rightly expressed concerns about access to commercial archives, and the ownership and selection of their contents (Mussell). The accuracy of optical character recognition (OCR) software has attracted much criticism, whilst other commentators have warned of the dangers of what might be termed ‘keyword blinkers’—arriving directly at a source and bypassing its wider context (Bingham 229–30). In particular, as John Plunkett has pointed out, the development of online archives has re-sensitised scholars to the materiality of the press (Plunkett 1). Digital archives are incapable of capturing the weight and texture of a newspaper page, whilst covers and other ephemera associated with the original text are often jettisoned. On a similar note, as Laurel Brake reminds us, digital archives have typically been created from microfilmed copies of bound volumes rather than original single issues (Brake). By the time we access them, many digital newspapers have been remediated three times (single issue > bound volume > microfilm > digitisation); each step serves to distance us from the original text. These concerns are not without foundation. When we digitise a newspaper, it is fundamentally changed. It is sensible, therefore, to ask what we lose in the process. But it is equally essential, and potentially far more rewarding, to ask the opposite question—to unpack the new, methodological opportunities created by the digitisation process and to explore what they mean for academic research.

Whilst the drawbacks of digital research have been repeatedly mapped out, the new opportunities offered by digital archives remain largely undefined. This is not to suggest a lack of enthusiasm for the topic. Rather, the advantages of digitisation have been treated as too obvious to require explanation. Researchers have been quick to grasp the practical benefits. Improvements in speed, access, volume and convenience are routinely celebrated. When asked to describe how digital archives have changed their lives, many historians highlight the fact that they no longer have to visit the British Library each time they want to consult a newspaper. Others rejoice that their lives are no longer blighted by malfunctioning microfilm readers. Keyword search engines are widely recognised as a time-saving device; a handy tool which helps researchers to find material quicker than by hand. So far, in other words, the mainstream profession has treated digitisation largely as a practical revolution—it has made research faster, easier, more convenient and more productive.

The effect of digitisation on our research has typically been understood as a by-product of these practical changes. As a result, discussions on the topic have tended to downplay the agency of researchers. Conference announcements regularly invite delegates to explore ‘how digitisation is impacting upon academic research’. This is not a bad question; there are important discussions to be had about the increasing academic interest in press history, about how other disciplines are using (or misusing) digital newspapers and about the democratisation of access to archives both in the academy and beyond. However, by focusing on this particular question, researchers have often ended
up considering ‘what digitisation is doing to historical research’, as if it was an uncontrollable force. To borrow from the nomenclature of digital publishing, this perspective has placed agency almost entirely with ‘content providers’ rather than ‘end users’. For much of the last 10 years, debates on digitisation have primarily been driven by librarians, digital publishers and a small community of digital humanists whose pioneering work has only recently begun to substantively engage the mainstream profession. As a result, the methodological implications of digital archives—their ability to extend the boundaries of research and answer questions that were previously unanswerable—have remained frustratingly absent from the debate.

Methodological discussions have, in part, been hamstrung by the speed of technological change. As Chris Willis pointed out in 2002, a website or digitisation project could change or collapse between the completion of an article and its publication (Willis 298). Similarly, the rapid advance of technology and the launch of increasingly sophisticated archives meant that methodological discussions always ran the risk of appearing dated. Whilst there is still no absolute guarantee of stability, we have reached a point where one of the core features of digital newspaper archives—the ability to perform basic keyword searches across multiple titles—is unlikely to change. Whilst the number of digitised titles will increase, the methodological implications of an archive containing 100 newspapers are largely the same as one with 1000. Moreover, new search technologies like text mining will be used alongside keyword searching rather than render it obsolete. There can be little doubt that the ability to search large archives of print culture for individual words or phrases will be a central feature of digital research for years to come. Methodologies developed for the archives of today will, therefore, not be rendered obsolete by the technology of tomorrow. There is no reason to defer our explorations any longer.

In recent years, there has been a growing recognition that digital methodologies require closer attention. In 2008, the Journal of Victorian Culture published a round-table discussion in which several contributors stressed the need to break free of established methodologies and develop new ways to approach digital newspaper archives. Jim Mussell argued:

> without developing...methodological approaches in how to think about and use [digital] resources, we remain trapped in methodologies shaped by our encounters with certain forms of printed objects in certain dusty rooms. (94)

Helen Rogers reached similar conclusions and argued that methodological issues now need to ‘be taken up by the users [of digital resources], working not only in web environments, but in the classroom, in essays and articles, and in academic conferences’ (Rogers 57). This is starting to happen. In 2009, the Research Society for Victorian Periodicals [RSVP], in partnership with Gale, launched the ‘Gale Dissertation Research Fellowship in Nineteenth-Century Media’. This annual award supports postgraduate research which makes innovative use of digital newspaper archives and is now in its fourth year. In 2010, a major conference on digital newspaper research was held at the British Library. In the same year, the University of Sheffield launched its ‘Centre for the Study of Journalism and History’, announcing that it was ‘particularly interested in developing robust methodologies for exploiting digital archives of journalism content’.
Most recently, an AHRC-sponsored research network entitled ‘Exploring the language of the popular in Anglo-American Newspapers, 1833–1988’ has placed digital methodologies at the heart of its agenda. The first meeting, held in January 2011, brought together scholars from a variety of countries and disciplines to discuss how online newspaper archives allow them to do new kinds of research.3

These early exchanges are beginning to stimulate a new, research-driven debate on digitisation. Whilst the conversation is still embryonic, attention is gradually coalescing around one key question: ‘what can we do that we couldn’t do before?’ Tom Scheinfeldt describes this as the ‘where’s the beef?’ question. Whilst he argues that users still need ‘time to play’ before putting digital tools to work, he recognises that the value of these resources will ultimately be measured by the new research they produce (Scheinfeldt). If media historians continue to produce the same kind of research as a decade ago, but simply use online archives to do it faster, then digitisation will have delivered a practical, but not an intellectual, revolution. At the time of writing, a formal body of scholarship driven by digital methodologies has yet to materialise. However, the increasing attention given to the subject suggests that its appearance is merely a matter of time. It is always dangerous to make predictions—particularly about technology—but over the course of the next decade, digital methodologies promise to have a significant impact on multiple areas of historical research. The scale of these developments could be comparable to the ‘cultural turn’ of the late 1980s when historians began to envisage and write new kinds of history based on a growing awareness of the power of culture and language in shaping past experiences and societies. In the process, they asked new questions, made new connections and developed new methodologies. We are potentially on the cusp of a similar revolution—a ‘digital turn’ in humanities scholarship driven by the creative use of online archives and a willingness to imagine new kinds of research.

These changes will take different forms in each field. Genealogists and biographers, for example, can use keyword searches to track down references to individuals. Literary critics can examine how a novel sat within wider cultural discourse. For linguists, digital newspapers offer extraordinary new opportunities to track the usage, circulation and evolution of language. Cultural history—a discipline which has traditionally distanced itself from computer-based research—arguably has the most to gain from digital methodologies. A central feature of cultural history has been its focus on the close examination of written and spoken texts, on the power of representation and place of discourse. An exploration of these discursive formations has fostered some valuable interdisciplinary work, but tracking cultural ideas across ‘a vast terra incognita of print’ (Leary Digital Age 206), whilst valuable and productive of exciting new insights, was severely limited in the pre-digital age. There were only so many texts that any scholar could be expected to study. Whilst this problem of volume has not been entirely resolved, the development of keyword search technology has made it possible to trace the development and movement of ideas and discursive formations in ways that were once impossible. In other words, this is cultural history, but not as we know it. The emphasis upon the importance of language and discourse remains, but the analytical tools at our disposal have been sharpened and expanded—an upgrade to what might be termed ‘Cultural History 2.0’.

The field of media history is an ideal space to pioneer these new methodologies. If there is to be a ‘digital turn’ in humanities scholarship, then the press will undoubtedly
play a central role. Whilst the digitisation of other textual sources like books, theatre scripts and letters is deserving of attention, newspapers, magazines and periodicals constitute the vast majority of newly digitised material. Moreover, the unique periodicity of the press makes it particularly well suited for studying continuity and change. As these sources become more central to our research, dialogues between media historians and scholars from other areas of the humanities will become increasingly important. For now, however, it also falls to media historians to help perform some of the theoretical groundwork for these emerging debates. In order to develop robust methodological responses to digital archives, we need to understand how newspapers are changed by digitisation. The following section offers some preliminary thoughts on how we might theorise this process.

**Deconstructing the Digital Newspaper**

The concept of a ‘digital methodology’ rests on one key idea: a hard copy of a newspaper is fundamentally different from a digitised version. At first glance, this difference seems obvious; one source is made from paper, the other exists as billions of 1s and 0s. However, the transformative effect of digitisation stretches beyond this material transition. Unlike microfilming, the creation of a digital newspaper does not simply produce what archivists term a ‘surrogate’, or stand-in, for the original. Instead, it creates something new; sources are ‘remediated’ and not just reproduced. Though a digitised text may look familiar, it is not the same source; we are able to access, read, organise and analyse it in radical new ways.

The nature of this transformation depends largely on the individual digitisation process. Most digital newspaper archives are currently designed around a relatively homogenous blueprint. They allow researchers to simultaneously search the full text of multiple periodicals, filter these searches by date, genre, title and relevancy, and then display the results as a list of images. Most allow users to construct more complex searches using ‘logical’ and ‘proximity operators’, ‘wildcards’ and ‘fuzzy searches’, and also make it possible to browse through individual issues. This formula has now become so ubiquitous that researchers would be forgiven for thinking that this is the only way to digitise newspaper archives. However, as anyone involved in a digitisation project will attest, developing these archives is a complex process—one made up of several stages, each of which requires a team of people to make difficult decisions. Choices have to be made about the selection of titles, the format of sources, the design of OCR and other bespoke image-processing software, and the manual correction of errors. These decisions have a substantial impact on the quality and quantity of data contained within the final digital source. The British Newspaper Archive, for example, allows users to manually correct OCR errors and tag articles with their own descriptive keywords. It is possible, for example, to tag articles with the term ‘Jack the Ripper’ even when the phrase does not appear in the original text. In time, this process of ‘crowd-sourced’ digitisation promises to improve the accuracy of the archive’s search engine and to create a new field within the data-set that supplements computerised text recognition with a more nuanced, human reading of the database’s contents.
The design of an archive’s search interface is equally central to determining the kind of questions that researchers can ask of its data. A change in interface can transform the methodological possibilities of an archive. The standard interface for Google Books, for example, allows researchers to keyword search the contents of millions of printed texts and then display them in Google’s standard list format. Recently, however, Google have developed a new interface named ‘Ngram Viewer’, a search engine which allows users to track the frequency of words and phrases over time and display results in graph form. Figure 1, for example, shows the results for the word ‘skedaddle’, an American slang term that entered British discourse in the mid-nineteenth century. This new interface draws upon the same raw data from Google’s book scanning project, but uses a new interface to interrogate the data in different ways. Similarly, a new project named Locating London’s Past allows users to pull information from a range of the existing data-sets (including court reports and coroners records) and plot it onto a map of the capital. Both projects demonstrate that digitised data need not be tied permanently to one interface or usage pattern. In theory, it would be possible to develop similar interfaces for the exploration of large-scale digitised newspaper archives. Gale’s forthcoming Nineteenth Century Collections Online will allow users to visualise keyword search results in graph form, though at the time of writing it is unclear whether the company has taken steps to ensure that fluctuations in the volume of searchable material do not distort comparisons between different periods.

By the time we view a digital archive, its contents have been through a complex process of transformation; the questions we can ask of a digital source are, in this sense, determined before we even login to the database. Unfortunately, this process of transformation is rarely shaped by academic researchers. Key decisions on the selection of an archive’s contents and the design of its interface are made by librarians, funding bodies and digitisation companies. When researchers finally enter the debate, it is often to complain about the flaws already embedded in the product. It would be easy to blame digitisation companies for this breakdown in communications but, in truth, the fault lies largely with us. We have, for the most part, failed to collectively identify and articulate our methodological requirements. In order for researchers to make the most of digitisation, it

---

**FIGURE 1**

A search for the term ‘skedaddle’ between 1840 and 1900 using Google’s Ngram Viewer
is essential that we become more active players in this design process and steer the development of new archives according to our research needs.

In order for these conversations to take place, basic methodological foundations need to be established. In particular, we need to explore the ways in which a digitised newspaper differs from its original paper. One of the key differences centres on the way in which information in these sources is organised and accessed. Classically, as Figure 2 shows, when we want to find something in a conventional newspaper we tend to approach it using a top-down approach. We start by selecting what we hope is the correct publication and then pick out what we hope will be the correct volume or issue. Next, we look at headers for potentially useful columns or articles. Finally, we read through the text of what promises to be a relevant article. At each stage, we exclude large quantities of information and hope that, through perseverance and experience, we will eventually unearth the information we seek. This is perfectly feasible for some themes; in order to explore press coverage of the Titanic disaster, it makes sense to select a range of papers from April 1912. But what if we want to look at the afterlife of the event, how it endured in British culture, how people dropped it into everyday conversation, how it became a metaphor applied to a variety of situations or began cropping up in jokes? Searching for these residual, and often deeply embedded, references using a top-down approach would be impractical, if not impossible—a case of looking for the proverbial needle in the haystack. The vast majority of the information in an article—the things that we can learn about the people and the society who produced and read it—are not organised using a top-down system. Once you get down to the article level in a newspaper, organisation breaks down and it becomes a complex mix of words and ideas. This is everyday cultural discourse—the fabric of everyday life in all its chaotic splendour. Like the ocean, it is teeming with life and with the potential for new discoveries; dip our head under the surface and we invariably find something new and interesting. But locating specific information using this top-down approach is often little more than a shot in the dark. With luck, time and effort we might indeed find useful data, but it is time consuming and unpredictable.

FIGURE 2
A conventional ‘top-down’ search for information in a printed newspaper archive
The digitisation of newspapers turns this established organisational framework upside down (Figure 3). Thanks to OCR technology and keyword search engines, we can now access this bottom level of text directly and navigate Leary’s ‘vast terra incognita of print’ with remarkable precision. Each individual word or phrase (provided it has been captured accurately by OCR software) is immediately detectable using the search engine. A new level of everyday cultural discourse is opened up. Billions of individual words—the fundamental building blocks of culture—are now at our fingertips. The search for information in newspapers need no longer be driven solely by top-down reading, but by the generation of new keywords. As Bingham argues, this process is by no means an easy uncritical procedure; finding the right keywords is often a lengthy business which requires perseverance and creativity, as well as an intimate knowledge of the culture and texts that are being studied (Bingham 229–30). Nor is to suggest that keyword searching replaces the need for extensive reading; if anything, by reducing the time spent searching through irrelevant articles, digital resources increase opportunities for close textual analysis (Bingham 229). Crucially, however, it is our direct access to individual words that presents the most exciting new methodological possibilities. Using the relatively simple interfaces already offered by commercial archives, researchers from across the humanities can now locate information, track the use and evolution of language, trace the development and movement of people, texts and ideas and explore discursive formations in ways that were hitherto impossible.

Applying Digital Methodologies

The most effective way to explore the methodological potential of online archives is to demonstrate its application. My doctoral research explored the role played by newspapers and periodicals in shaping the Victorian public’s relationship with America (Nicholson Looming Large). It was devised in response to the launch of the 19th Century British Library Newspaper Archive in 2007 and was among the first history projects to be

FIGURE 3
A ‘bottom-up’ keyword search for information in a digital newspaper archive
designed specifically around the new possibilities of newspaper databases. In order to track the changing presence of America in British newspapers, it was necessary to develop two contrasting digital methodologies. The first—what we might term a ‘macro’ approach—used quantitative methodologies to identify broad patterns in the database. A full text search for the term ‘America’ in the 37-year period covered by my study returns 881,900 hits. To this number we must add the results for related keyword searches, such as United States (743,922), New York (1,060,181), Chicago (272,942) and Yankee (35,984). Even the most dedicated scholar would baulk at the prospect of reading each of these articles. And yet, to restrict them to a manageable sample would be to eliminate more than 95% of the archive’s contents. The solution to this problem was to practice a form of what the digital humanist and literary scholar, Franco Moretti, terms ‘distant reading’ (Moretti).6 In this case, close reading was replaced by computer-aided counting. I carefully selected a representative sample of 29 papers which covered the whole of my period without substantial missing issues, devised a complex series of keyword searches designed to identify articles about America and exclude irrelevant texts and then tracked the number of hits over time.

The results make for interesting reading. Figure 4 displays the results of a search for ‘Americ* OR United States’. As we might expect, references to the country increased dramatically during the early 1860s as British papers reported on the progress of the civil war. In the late 1860s, as the conflict gradually receded from view, references to America experienced a slight decline. They increased steadily during the 1870s before spiking

![Figure 4](image)

**FIGURE 4**
Full-text search for ‘Americ*’ OR ‘United States’ in 29 British newspapers (results displayed by year)
dramatically between the years 1879 and 1881. In 1881, references to America were 41% higher than they had been in 1878. Whilst it would be unwise to place too much interpretative weight on such crudely derived data, a qualitative study of these years revealed a range of significant (and hitherto unexplored) Anglo-American debates, cultural encounters and technological developments. One of the factors that contributed to the US’ increased media presence in this period was a growing concern in Britain about transatlantic economic competition. Figure 5 shows the results of a search for all articles in which references to America, Germany and France appeared within 10 words of the term ‘competition’. As we can see, articles linking ‘America’ and ‘competition’ spiked dramatically in the early 1880s. It was then possible to focus in on this particular period, subject it to a closer qualitative analysis and identify the factors that contributed to this media attention. It was also possible to conduct a more focused qualitative search. Figure 6 shows results broken down by month, and reveals what appears to be a seasonal pattern; concerns about American competition appear to have increased in the autumn months, possibly in response to the harvest of British crops. A team of Harvard scientists has recently dubbed this form of quantitative research ‘culuromics’ (Michel). Whilst it is important to recognise the limitations of such an approach—it does not, after all, reveal the meaning of the texts it counts—it provides a useful way to visualise broad cultural trends and identify areas for closer exploration.

The second digital methodology employed in my project adopted what might be termed a ‘micro’ approach. Rather than identify broad trends, it used keyword searches to

![Figure 5](image)

**FIGURE 5**
Articles in which America, Germany and France are mentioned within 10 words of ‘competition’ (results displayed by year)
track the transatlantic movement of specific words and texts. One of the key aims of the project was to look at how the Victorians encountered, discussed and used American English—including ‘Americanisms’ and American slang. My first step was to put terms such as ‘Americanism’, ‘American English’ and ‘Yankee Slang’ into the search engine. Predictably, I was overwhelmed by material—the word ‘Americanism’, for example, appears 1864 times in the 19th Century British Library Newspaper archive alone. Before wading through this material, I modified the search so that it would focus purely on article titles. This left me with 42 articles, all of which seemed to be discussing the subject of ‘Americanisms’. Some of these articles were written by journalists, but most were extracts clipped from essays in monthly reviews. Many of these publications are currently available online, so it was possible to trace an extract back to its original source. I could then examine the source, see which part was extracted, and search to see if other newspapers had cut and pasted different sections. However, I did not want to confine my analysis to the way in which literary critics or philologists discussed Americanisms, but wanted to explore how these words were used by British people and how they circulated in everyday cultural discourse.

Using John Russell Bartlett’s Dictionary of Americanisms (1848), my next step was to enter specific American words and phrases in the search engine. Simply by looking at the number of hits, it was possible to gain some sense of whether a particular Americanism circulated extensively in Britain or met with little success. So, for example, the phrase ‘eyes peeled’ (as in, ‘keep your eyes peeled’) only returned 23 hits on the British Library database, mostly reprints of the same advert. ‘Eye opener’, on the other hand, returned 927 hits, suggesting that it circulated more widely. On examining the articles, I was able to access the complex contexts in which a word or phrase might be used—to see who used it, as well as where, when and how it was used. I could use the press as a living dictionary, one which allowed me to access the minutiae of everyday cultural discourse and track the complex ways in which words and phrases circulated, how their meanings evolved and
how their usage changed on an almost daily basis. It was possible, for example, to track how the word ‘skedaddle’ arrived in Britain in the early 1860s as part of press reports on the American Civil War and went on to take the country by storm. Letters debating its etymology were sent to The Times and were reprinted in several provincial papers; a Glasgow-based merchant used it in advertisements for rocking chairs; a boat named ‘Skedaddle’ took part in the Durham and Londonderry Regattas; several racehorses were named after it; a game named ‘skedaddle’ became popular with children; an energetic dance named ‘The Skedaddle Breakdown’ was performed nightly at the Haymarket; Punch and other comic papers quickly inserted it into the dialogue of their American characters; many newspapers and periodicals printed it in a range of different contexts; readers throughout the country responded by using it in letters to editors. This was just the tip of the iceberg—the word continued to circulate in venues ranging from racecourses to parliament, long after the initial waves of ‘skedaddlmania’ subsided. Without keyword searching, the cultural evolution of this word—its changing meanings and connotations, the differing contexts in which it was used and the dynamic way it moved around a transatlantic cultural circuit—could never have been tracked.7

Crucially, digital archives allowed me to map the role played by the press in this process. Newspapers and periodicals were crucial to the importation of American culture and its subsequent circulation throughout Britain. The latest American slang, for example, often arrived in the country as part of jokes clipped by British editors from American newspapers. As digital archives become increasingly accessible, a growing number of researchers are using keyword searches to cherry-pick material from newspapers with little regard to context. Moreover, as linguists and cultural historians begin to apply digital methodologies to large-scale press archives, there is a danger that newspapers will be treated as mere buckets of words with no agency or function of their own. In the coming years, one of the biggest challenges facing press historians will be to ensure that the historical agency and complex materiality of newspapers are not forgotten in a rush to mine their contents. Fortunately, digital archives present ample opportunities for exploring the historical significance of the press. Whilst my own research started life as little more than a plundering expedition in which I hoped to mine press archives for Victorian public opinion, it has become increasingly clear to me that newspapers played a key role in shaping the nineteenth-century Anglo-American relationship, and that this process was an important field of enquiry in itself. Newspapers did not simply reflect British responses to the USA, but were central to their formation. Without digital archives, it would have been impossible to map out the complex networks and connections that made up this transatlantic cultural relationship.

Conclusion

Digitisation has allowed me to undertake research that was previously impossible—to ask new questions, make new connections and map out a relatively unexplored area of transatlantic popular culture. However, the techniques discussed in this article represent only an early experimentation with digital methodologies. In the coming years, it is likely that researchers from a range of disciplines will develop more creative and nuanced approaches. It is difficult to discern whether this research will spark a dramatic ‘digital turn’
in humanities research, or whether they will creep imperceptibly into the fabric of everyday scholarship. One thing is clear. Whether we seek to uncover the life of a particular individual, explore the evolution of language, track the development and circulation of an idea, or write the history of an individual publication, keyword-searchable newspaper databases will allow us to work in new ways. It is crucial that researchers—and media historians in particular—start to unpack these methodological possibilities, to find out what we can, and cannot, do with digitised sources, to share our ideas in structured discussions and take a more active role in the development of new archives and interfaces.

Acknowledgements

I am grateful to the AHRC and the Research Society for Victorian Periodicals for supporting this research.

Notes

1. For an extensive review of the Burney Collection, see Ashley Marshall and Robert D. Hume.
4. The potential of crowd-sourced digitisation has recently been demonstrated by the success of the Dickens Journals Online project, which has mobilised a team of volunteers to manually correct OCR errors in Household Words and All the Year Round. http://www.djo.org.uk/
6. For more discussion of Moretti, distant reading and the Victorian press, see Bob Nicholson ‘Counting Culture’.
7. A similar approach was used to track the transatlantic circulation of jokes. See Bob Nicholson, ‘You Kick the Bucket’.

References

MICHEL, JEAN-BAPTISTE, YUAN KUI SHEN, AVIVA PRESSER AIDEN, ADRIAN VERES, MATTHEW K. GRAY, THE GOOGLE BOOKS TEAM, JOSEPH P. PICKETT, DALE HOIBERG, DAN CLANCY,


Bob Nicholson, Department of English and History Edge Hill University, St Helens Road, Ormskirk, Lancashire L39 4QP, UK. Tel: +01695 650 992; E-mail: Bob.Nicholson@edgehill.ac.uk