Routledge
Taylor & Francis Group

Cultural and Social History
The Journal of the Social History Society

ISSN: 1478-0038 (Print) 1478-0046 (Online) Journal homepage: http://www.tandfonline.com/loi/rfcs20

I'd Rather be a Librarian

Andrew Prescott

To cite this article: Andrew Prescott (2014) I'd Rather be a Librarian, Cultural and Social History, 11:3, 335-341, DOI: 10.2752/147800414X13983595303192

To link to this article: http://dx.doi.org/10.2752/147800414X13983595303192


Published online: 01 May 2015.


Submit your article to this journal


Article views: 42


View related articles


View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=rfcs20

Download by: [Universite Laval] **Date:** 09 May 2016, At: 09:53

I'D RATHER BE A LIBRARIAN

A RESPONSE TO TIM HITCHCOCK, 'CONFRONTING THE DIGITAL'

Andrew Prescott
King's College London

Keywords: digital humanities, librarianship, canonicity

Downloaded by [Universite Laval] at 09:53 09 May 2016

In describing the Industrial Revolution, Sidney Pollard emphasized how some industries changed more quickly than others. Pollard described how a 'visitor to the metalworking areas of Birmingham or Sheffield in the mid-nineteenth century would have found little to distinguish them superficially from the same industries a hundred years earlier. The men worked as independent contractors in their own or rented workshops using their own or hired equipment … These industries … were still waiting for their Industrial Revolution.'[1] Yet the environment in which these men were working had been transformed. Their wheels were now powered by steam, and minor operations such as stamping and cutting had been speeded up by the use of machinery. The workshop might be lit by gas and have a direct water supply. Railways made distribution easier and cheaper while also giving access to a large labour market.

Perhaps academic historians today are in a similar position to those craftsmen of the mid-nineteenth century. The way in which we access books and journal articles, our methods of retrieving materials from libraries and archives, and the way in which our books and journal articles are published have all changed radically over the past twenty-five years. Yet, as Tim Hitchcock emphasizes, the fundamental nature of the way in which we think, discuss and write about the past does not appear to have changed very much. I nowadays use a word processor to perform more quickly the tasks that I laboriously undertook thirty years ago with pen, paper, typewriter and correcting fluid, but although the use of a word processor has probably affected my literary style, it doesn't fundamentally affect the way I build and express an historical argument. Am I like a Sheffield cutler of the nineteenth century, gratefully accepting the help of machines to make my life less arduous, but still awaiting my Industrial Revolution? Are conservatism and complacency preventing historians from seizing the chance to change the nature of historical discourse in a way that would build new audiences for history and create a stronger historical presence in public culture, as Tim Hitchcock powerfully argues?

I have spent twenty-three years of my professional life as a curator or librarian, and eleven years as an academic (although, despite having trained as a historian, never in a

Address for correspondence: Professor Andrew Prescott, Department of Digital Humanities, King's College London, London WC2R 2LS, UK. E-mail: andrewjprescott@gmail.com

Cultural and Social History, Volume 11, Issue 3, pp. 335–341 © The Social History Society 2014
DOI 10.2752/147800414X13983595303192

conventional history department). I have no doubt that the most intellectually exciting, stimulating and challenging years of my career have been my time as a librarian. By contrast, my spells as an academic have seemed intellectually moribund, stultified by the dead hand of academic inwardness and self-absorption, made more depressing by the pernicious influence of British bureaucratic concerns such as student surveys or the Research Excellence Framework (REF). The reason that life as a librarian has been much more exciting is because it has provided a wonderful front-row seat to witness the progress of that revolution in access to historical materials so vividly described by Tim Hitchcock. Everyone hopes to have been in the right time and the right place at some point in their life, and I feel that, in witnessing as a librarian during the 1990s the creation and growth of many digital tools and resources, I was experiencing at first hand major historical developments. I thank my lucky stars that I took a deliberate decision not to become a conventional academic historian in 1979 and went down a curatorial route instead.

When I first joined the British Library in 1979, it was evident that the most lively and vibrant unit was the young team working on the creation of the Eighteenth Century Short Title Catalogue (ESTC), under the charismatic leadership of the historical bibliographer Robin Alston. Alston had been insistent that the ESTC should from its inception be planned as a computerized resource, and his determination on this fundamental point has paid dividends for scholars. The Eighteenth Century Short Title Catalogue subsequently became the English Short Title Catalogue, extending its coverage to all material published in Britain, Ireland, the United States and the British Empire between 1473 and 1800, and providing for that period a comprehensive database of surviving publications in the English-speaking world. The ESTC was available online (initially as a charged service, now free) from an early date, and an obvious next step was to make the books listed in the catalogue available via microfilm. By these means, it was possible for the first time to access in one place the entire surviving archive of pre-1800 English language printing.

It was this pioneering work which paved the way for the development of the digital services with which we are now familiar, such as Early English Books Online (EEBO) and Eighteenth Century Collections Online (ECCO). This background is important to remember, since it reminds us that the roots of these packages lay in the process of identifying and recording for the first time the nature of the surviving print archive. Full text search was never a primary aim of these projects and indeed was not initially included in EEBO. Since full text search was a subsidiary aspect of these projects, it is not surprising that we encounter the sort of OCR problems described by Tim Hitchcock and which are a major issue for users of ECCO. History as a discipline is grounded in critical use of textual and other materials from the past, which means that – as Hitchcock rightly insists – historians have an inescapable duty to understand the way in which primary materials are presented to them and the way in which the limitations of this may distort their investigation of the past. This necessitates having some appreciation of the limitations of the OCR in ECCO, as well as grasping how the work of the Text Creation Partnership means that full text search issues are different for EEBO.

Downloaded by [Universite Laval] at 09:53 09 May 2016

Another project in which Robin Alston played a part while I was at the British Library was the digitization of the Burney Newspapers. As the major collection of eighteenth-century newspapers, the Burney Collection was in constant demand in the British Library Reading Room. To reduce wear and tear on the original newspapers, the Burney Collection was only made available to readers on microfilm. Not only is microfilm an ungrateful medium to use, but there was also only one microfilm of the Burney Newspapers available in the main reading room.[2] The British Library had recently acquired a microfilm scanner, and the Burney Newspapers were an obvious first candidate for experimentation. It was hoped both that the digital images would be easier to manipulate than microfilm, and also that it would be possible for several readers to consult the digital surrogate simultaneously, thereby eliminating a major reading room bottleneck. Again, search was not a major consideration. The scanning of the microfilms was successful, but the British Library lacked the resources to make the images widely available, and it was only some years later that these scans were developed in partnership with Gale-Cengage into the now familiar online version of the Burney Newspapers. The limitations of the OCR searching in the online Burney Newspapers reflect the origins of the project in seeking to provide more efficient access to reading room surrogates – the concern of the librarians involved was simply to make life easier in the reading room, and creating a searchable archive was not a major consideration.

Since the focus of the original microfilm scanning project was on improving reading room access to a high-use collection, other eighteenth-century newspapers were ignored. Whereas ECCO is based on the ESTC and thus aspires to comprehensive coverage of the eighteenth-century web archive, online coverage of eighteenth-century newspapers is based on a collection whose core consists of left-over papers which Burney managed to pick up from the tables of a coffee house run by his maiden aunts. Burney's collection is consequently chiefly focused on London, and coverage of provincial newspapers is patchy. Many runs are incomplete. Yet too many users seem to assume that this web resource is a comprehensive archive of British eighteenth-century newspapers. It isn't. As a result of the online prominence of the Burney newspapers and the ease with which the collection can be consulted, the biases of selection and coverage inherent in the Burney collection are being silently incorporated into wider historiography. Once the Burney collection has been in use for another ten years or so, it will be interesting to undertake a study to analyse patterns of citation of eighteenth-century newspapers. I bet my British Library Reader's ticket that we will find a decline in use of provincial newspapers and an overwhelming bias towards London material.

New canons of primary material are being quietly established without much influence from or critical involvement by academic historians. The fact that *The Times* archive is nationally available via an agreement with Gale-Cengage is reinforcing the prominence of *The Times* in historiography at the expense of other newspapers such as the *Manchester Guardian*, the *Observer* or the *Scotsman*, which are not so commonly accessible online. While the spade work done by the bibliographers who compiled the ESTC means that for pre-1800 materials our digital print archive is comprehensive,

Downloaded by [Universite Laval] at 09:53 09 May 2016

unfortunately after 1800 we do not have such bibliographies available. The Nineteenth Century Short Title Catalogue aspires to similar coverage to the ESTC for the period from 1801 to 1918, but this is a vast undertaking and is far from complete. The major book digitization projects that Tim Hitchcock refers to, such as most notably Google Books, have worked on the assumption that what are considered in Silicon Valley to be the world's greatest libraries contain all the world's knowledge – a breathtaking piece of Californian optimism. The 'library partners' of Google Books include all the usual suspects: Harvard University, Princeton University, Columbia University, University of Michigan, the University of California, Oxford University, the British Library – a sort of ivy-league view of human knowledge that excludes not only local libraries and archives but also what are seen from Mountain View as second-rank libraries, such as the National Library of Wales or the National Library of Scotland.

These digital canonicities go beyond the superficial techno-optimism that permeates Google Books. It would be reasonable to assume that the British Library catalogue service COPAC provides comprehensive coverage for the UK's major research libraries. COPAC declares that it offers access to 'a wide and varied range of library catalogues, from the collections of the Oxford and Cambridge universities to the libraries at the National Trust and the Royal Botanic Gardens Kew', and states that its contributors 'run the gamut from conservatoires to the catalogues of the major Russian and East European Collections in the country'. Another of my memorable digital moments as a librarian was spending a weekend in a hotel in York with a group of senior British librarians engaged in the tense negotiations which enabled the British Library's online catalogue to be linked to the union catalogue of the Consortium of University Research Libraries in order to facilitate the creation of COPAC. The Consortium of University Research Libraries (now Research Libraries UK) fundamentally consisted of the libraries of what has become the Russell Group of universities. So, RLUK includes the University of Sheffield, which has quite limited library special collections, but excludes such universities as Sussex and Aberystwyth which have some significant special collections. As a result, the collections at Aberystwyth and Sussex (and many other universities) do not appear in COPAC. They may appear in the more extensive Worldcat, but the content of Worldcat can also be shaped by administrative quirks.

As a result of these digital canonicities, books can disappear from the scholarly gaze. In her recent book *Liberty's Dawn: A People's History of the Industrial Revolution* (New Haven, CT, and London: Yale University Press, 2013), Emma Griffin has used autobiographies by working men and women to re-examine the debate about the effects of the Industrial Revolution on the standard of living and quality of life. Many of the autobiographies used by Griffin were listed in a monumental annotated critical bibliography of *The Autobiography of the Working Class* edited by John Burnett, David Vincent and David Mayall, published in three volumes from 1984 to 1989 and listing over two thousand autobiographies by people of working-class origin produced between 1790 and 1945. Many of the items listed in this bibliography survive only in a handful of copies and are often found in local libraries or archives – precisely the sort of institutions that Google Books assumes are marginal to the world's knowledge.

Downloaded by [Universite Laval] at 09:53 09 May 2016

A cross-comparison of the bibliography of *The Autobiography of the Working Class* with Google Books, COPAC and Worldcat is disconcerting. A significant number of items in the bibliography cannot be traced in any of the major online resources. For example, Benjamin North was born at Thame in Oxford in 1811, the eighth child of a labourer. He was a boy shepherd, bird-keeper, plough-boy and groom, then trained as a paper-maker, but he was made redundant by the introduction of new machinery. He eventually became a traveller for a chair-maker and set up a successful furniture business in High Wycombe. North's autobiography was published after his death by his son and is preserved in the local studies library at High Wycombe, but it isn't in COPAC, Worldcat or Google Books. North's voice has been effectively silenced. John Finney worked in the Potteries from the age of thirteen and in 1902 published *Sixty Years Recollection of an Etruscan*, a copy of which is preserved in the Horace Barks Reference Library in Hanley. Like many such local libraries, the catalogue at Hanley is still apparently a card catalogue, and the online catalogue for Stoke libraries does not refer to Finney's book. So, once again, it is absent from our major online catalogues, and Google Books denies all knowledge of such a book as *Sixty Years Recollection of an Etruscan*.[3]

By concentrating on larger libraries, Google Books silently hardwires many cultural and social biases into its presentation of the world's knowledge. Those writers of working-class origins who had a success story to report, who had become distinguished statesmen, successful businessmen or religious leaders, were able to find commercial publishers. Writers whose life demonstrated the virtues of temperance, prudence and self-help were particularly favoured. Books published by major commercial publishers would be collected by legal deposit libraries in Britain, and might even excite interest across the Atlantic. As a result, it is these volumes which we tend to find in such resources as Google Books; those which were locally published and are only found in local repositories are ignored by Google. Surely, it should be an urgent mission for all historians to investigate the extent and nature of these sorts of biases.

The most serious aspect of the emergence of new digital canons for academic history is the effect of subscription-based family history websites. While firms like Ancestry and D.C. Thomson have shown great enterprise in digitizing large parts of the historical archive for use by family historians, the effect has frequently been to fragment the availability and use of major categories of historical records. This is because family history is big business, with one analyst projecting that the market will be worth $4.3 billion by 2018.[4] According to *BusinessWeek*, 'genealogy ranks second only to porn as the most searched topic online'.[5] This means that firms like Thomson have the resources to undertake large-scale digitization for archives and libraries, but they do not have any incentive to make them available for university use through collective licensing schemes such as those run by JISC (formerly the Joint Information Systems Committee). The baleful effect of this situation is evident from the British Newspaper Archive, developed by Thomson with the British Library, which contains over seven million images of newspapers from the British Library's collections. The British Newspaper Archive does not have a licensing agreement with JISC and is not widely available through British university libraries. Historians wishing to use this collection

Downloaded by [Universite Laval] at 09:53 09 May 2016

generally need to take out an annual subscription which costs £79. By contrast, the collection of nineteenth-century British newspapers published by Gale-Cengage, which is smaller and has a different selection of material, is widely available to British academics as a result of a licensing arrangement with JISC. Consequently, the Gale-Cengage package is more likely to be used by historians than the British Newspaper Archive, which has much fuller coverage of local newspapers. What is the effect of this likely to be on historiography? Nobody knows.

Archivists have been very enthusiastic about making agreements with firms such as D.C. Thomson, but this effective privatization of much of the British historical archive, on a financial model which a distinguished practitioner of digital humanities has compared to pay-day loans, has passed largely undiscussed by the historical profession. Rate books have long been considered one of the mainstays of local historical studies in England. Ratebooks for Westminster, Southwark, Manchester and Plymouth are now available online through Thomson's 'findmypast' service, but any historians wishing to access them will not only have to grapple with an interface designed exclusively for family historians but also have to pay at least £100 a year to do so. The scale of digitization that has been undertaken for family historians in this way is huge but (apart from the censuses) doesn't seem to have been noticed by many academic historians. In the past month alone, it has been announced that the British Library has recently released over 2.5 million genealogical records from the India Office records via 'findmypast', while the valuation rolls from Scotland, another key source for local studies, are also being made available on a subscription basis.

Surely, any historian worth their salt would want to get passionately involved in debate and discussion about how these fundamental historical records are made available and how we are able to interrogate them. Tim Hitchcock articulates how all historians should feel, confronted with such a major shift in the nature of our engagement with the past. But the only time that the corduroyed Colonel Blimps of the British historical establishment have grudgingly bestirred themselves from their deep slumber to engage extensively with digital matters was when they belatedly realized that changes in open access might upset the cosy financial arrangements that provided a life-support system for ailing learned societies, and a hasty rearguard action was mounted to try to preserve the status quo.

It can hardly be surprising that, wanting to contribute to our engagement with these changes, I have found that librarianship provides a better route to creating new intellectual visions and embracing wider audiences than those self-regarding and incestuous debates that too often pass for historical scholarship. I support Tim in his clarion call to academic historians to confront the challenge of the digital, but unfortunately I feel that the historical profession has become so internalized and cut off from wider intellectual cultures that it will prove to be beyond redemption, and we may need to look elsewhere to find the response that Tim seeks. Perhaps, dare I say it, it is the librarians and archivists who are already rising to this challenge and who will ensure that history has its Industrial Revolution.

Downloaded by [Universite Laval] at 09:53 09 May 2016

NOTES

1. Sidney Pollard, *Peaceful Conquest: The Industrialisation of Europe 1760–1970* (Oxford: Oxford University Press, 1981), p. 27.
2. This microfilm set now sadly sits unused on the shelves of the British Library's Rare Books and Music Reading Room at St Pancras.
3. I have discussed these issues about working class autobiography more fully on my blog: http://digitalriffs.blogspot.co.uk/2013/08/how-web-can-makes-books-vanish.html.
4. Mike Richman, 'Technology, Word of Mouth Help Genealogy Hit the Mainstream', *Voice of America*, 26 September 2013: http://www.voanews.com/content/technology-and-word-of-mouth-help-genealogy-hit-mainstream/1757389.html (accessed 13 February 2013).
5. Bruce Falconer, 'Ancestry.com's Genealogical Juggernaut', *Businessweek*, 20 September 2012: http://www.businessweek.com/articles/2012-09-20/ancestry-dot-coms-genealogical-juggernaut (accessed 13 February 2014).

Downloaded by [Universite Laval] at 09:53 09 May 2016