Transactions of the RHS 25 (2015), pp. 75–93 © Royal Historical Society 2015
doi:10.1017/S0080440115000031

MAKING HISTORY ONLINE

The Colin Matthews Lecture for the Public Understanding of History

By Tim Hitchcock and Robert Shoemaker

READ 12 NOVEMBER 2014

ABSTRACT. This article considers the implications of recent innovations in digital history for the relationship between the academy and the public. It argues that while digitisation and the internet have attracted large new audiences, academic historians have been reluctant to engage with this new public. We suggest that recent innovations in academic digital history, such as the highly technocratic 'Culturomics' movement, have had the unintended effect of driving a wedge between higher education and the wider public. Similarly, academic history writing has been slow to embrace the possibilities of the internet as a means of dissemination and engagement; and academic publishing has moved even more reluctantly. Despite these issues, this article argues that the internet offers real opportunities for bridging the divide between the academy and a wider audience. Through non-traditional forms of publication such as blogging; through Open Access policies; and through new forms of visualisation of complex data, the digital and online allow us to present complex history to a wider audience. We conclude that historians need to embrace the 'affordances' and 'disruptions' posed by the internet to render the discipline more open and democratically accessible.

In the last twenty years the way in which we do historical research has been transformed by digitisation and the internet. With JSTOR, which has provided online access to a vast archive of periodical articles since 1995, Early English Books and Eighteenth-Century Collections Online (EEBO and ECCO, from 1998), Google Books (2004) and numerous other online resources, the traditional journey into the library and the archive and back out again has been reshaped.[1] We might call this the creation of the Western print archive – second edition.

Through this vast array of electronic resources, emerging from the private sector, museums and archives and higher education, British history in particular has been made newly available at the click of a mouse to anyone with an internet connection. Of course, there are limitations – issues of coverage and what is left undigitised, of access and paywalls, and of OCR quality and copyright – but in less than a generation, the British

[1] See www.jstor.org; www.eebo.chadwyck.com; www.gale.cengage.co.uk/product-highlights/history/eighteenth-century-collections-online.aspx; http://books.google.com.

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

past, particularly prior to the twentieth century when copyright restricts access, has become the most digitised where and when in the world.

We have been hugely privileged to have been able to contribute to this phenomenon, through our involvement in the creation of the Old Bailey Online, launched in 2003, London Lives and Locating London's Past in 2010, Connected Histories in 2011 and the ongoing Digital Panopticon project.[2] This experience has given us a strong sense both of the possibilities of digital history and its risks. A fundamental feature of digital history is the way it opens up history to the public, but the very publicity of the internet has proved problematic for academic historians. In this paper, we consider how recent innovations in the online provision of resources, and in the crowdsourcing and co-creation of research materials, have the potential to reconfigure the relationship between the academy and the public; and argue that for this to happen, academic historians need to embrace these new opportunities.

Digital resources on the internet have let historians do remarkable things. Just by way of a simple example, take the Londoner Sarah Durrant, one of almost 125,000 defendants convicted at the Old Bailey in the nineteenth century. Her experience was by no means unique, but quite suddenly that experience is available to us in a new way. Sarah claimed to have found two bank notes on the floor of the coffee house she ran in the London Road in 1871, at which point she pocketed them. In fact they had been stolen from the briefcase of Sydney Tomlin, at the Birkbeck Bank, Chancery Lane, a few days earlier. From her prison records, we can learn about her widowed status and physical description, including the existence of two moles on her face – one on her nose and the other on her chin (Figure 1). We can see her scared and resentful eyes staring at us from a mug shot. We also have the words recorded at her trial, from which we know that Sarah was convicted of receiving stolen goods; and that she had been turned in by a Mrs Seyfert – a drunk, to whom Durrant had refused a hand-out (Figure 2). On 11 January, she was found guilty of receiving stolen goods and sentenced to two years of imprisonment. And we can read the newspaper report of the same trial (Figure 3).

Later that year, while she was in Wandsworth Gaol, she was recorded in the census, along with all her fellow prisoners (Figure 4). From her trial, we know where Durrant had been living when the crime took place: in Southwark, at No. 1 London Road. We know that she was a little uncertain about her age from the different answers she provided to different clerks, and from the census we can find out both who she shared prison life with following her conviction, and who lived up one flight of stairs, and down

[2] See www.oldbaileyonline.org; www.londonlives.org; www.locatinglondon.org; www.connectedhistories.org; www.digitalpanopticon.org.

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

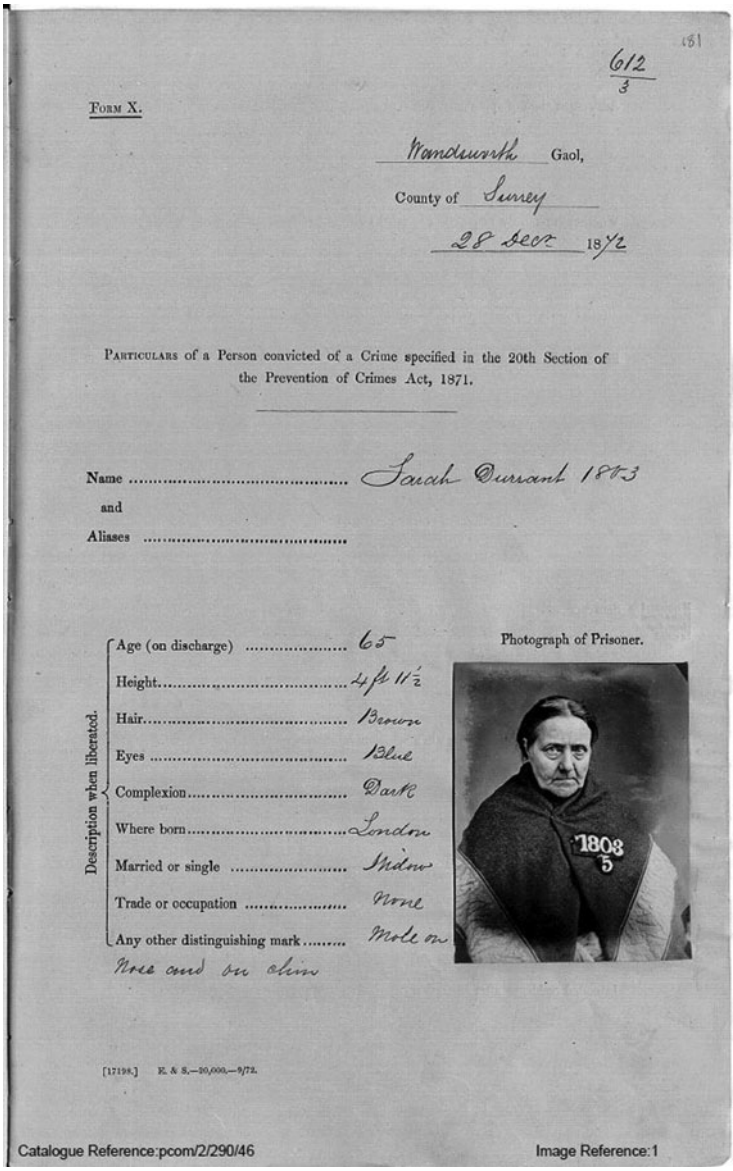


Figure 1 The National Archives: 'Prisoners Photograph Album', Sarah Durrant, 28 December 1872, PCOM2/290/46 (http://discovery.nationalarchives.gov.uk/details/r/C9072200, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press



Figure 2 (Colour online) *Old Bailey Proceedings Online* (www.oldbaileyonline.org, version 7.2, 28 Apr. 2015), January 1871, trial of Sarah Durrant (t18710109–137).

another in her previous home in Southwark. From here, it is a small step on the web to go to the Charles Booth Online Archive posted by the London School of Economics in 2001 (a digitised version of Booth's *Life and Labour of the People in London*, 1889–1903), which in turn lets us know a bit more about the street and its residents (Figure 5). According to a policeman's notebook and Booth's map, the street was 'a busy shopping street', with the social class of the residents declining sharply to the west, where it comprised 'some comfortable [households], others poor'.[3]

In half an hour's search we can put together a life, an experience and an emotional and empathetic contact with one of the more than three million mostly anonymous (to us) men and women who lived in London in 1871.

In the process of making this kind of research possible, digitisation and the internet have helped spirit into being new audiences for history, new practitioners of history writing and new forms of historical practice. That there was an enormous appetite for resources like this was made strikingly

[3] *Charles Booth Online Archive*: Police Notebooks, 'District 31: Lambeth and St Saviour's Southwark', B363, pp. 16–17 (http://booth.lse.ac.uk/static/b/districts.html, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

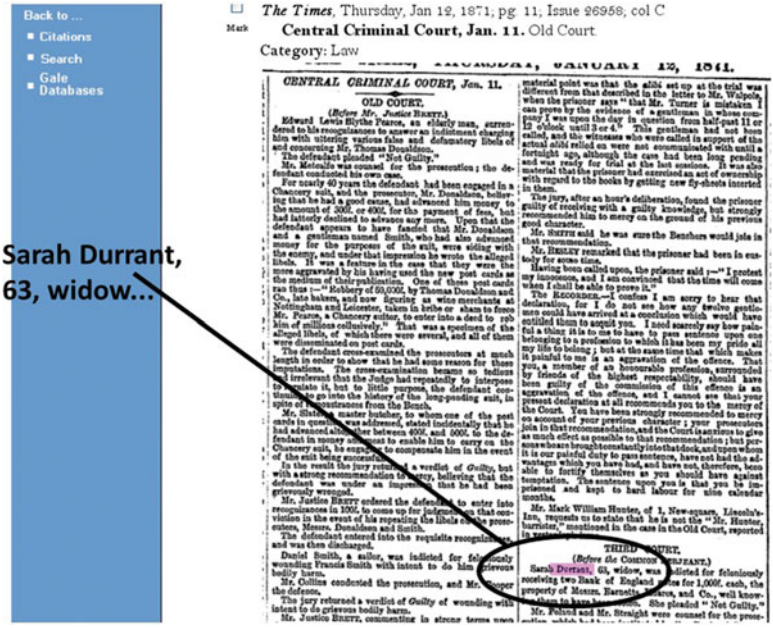
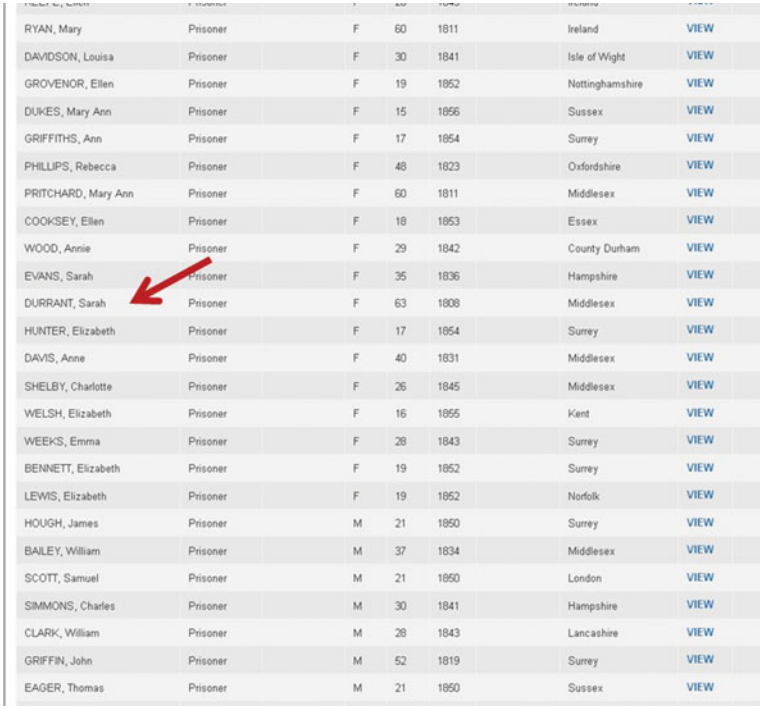


Figure 3 (Colour online) *The Times*, Thursday, 12 January 1871, 11; issue 26958 (http://gale.cengage.co.uk/times.aspx/, 28 Apr. 2015).

apparent in January 2002, when the first British census was made available online – and immediately crashed through overwhelming demand.[4] Over the past fifteen years, following the creation of dozens of similar resources, it has become clear that a new cohort of historians has come into being. Despite the fact some online resources depend on institutionally based subscription models which effectively exclude independent scholars, a vast amount of material online is readily accessible to everyone. The public sector, which posted all the sources of information about Sarah Durrant cited above except the newspaper article, typically provides free online access. While privately run resources such as Ancestry and Findmypast charge their users, most have developed business models based on affordable subscriptions, while others, such as Google Books, are free. As a result, the non-academic historian with internet access has at their fingertips more real data than can be found in any single archive or hard-copy library.

[4] *BBC News*, 'Census website a crashing success', 2 Jan. 2002 (http://news.bbc.co.uk/1/hi/uk/1737861.stm, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press



RYAN, Mary	Prisoner	F	60	1811	Ireland	VIEW
DAVIDSON, Louisa	Prisoner	F	30	1841	Isle of Wight	VIEW
GROVENOR, Ellen	Prisoner	F	19	1852	Nottinghamshire	VIEW
DUKES, Mary Ann	Prisoner	F	15	1856	Sussex	VIEW
GRIFFITHS, Ann	Prisoner	F	17	1854	Surrey	VIEW
PHILLIPS, Rebecca	Prisoner	F	48	1823	Oxfordshire	VIEW
PRITCHARD, Mary Ann	Prisoner	F	60	1811	Middlesex	VIEW
COOKSEY, Ellen	Prisoner	F	18	1853	Essex	VIEW
WOOD, Annie	Prisoner	F	29	1842	County Durham	VIEW
EVANS, Sarah	Prisoner	F	35	1836	Hampshire	VIEW
DURRANT, Sarah	Prisoner	F	63	1808	Middlesex	VIEW
HUNTER, Elizabeth	Prisoner	F	17	1854	Surrey	VIEW
DAVIS, Anne	Prisoner	F	40	1831	Middlesex	VIEW
SHELBY, Charlotte	Prisoner	F	26	1845	Middlesex	VIEW
WELSH, Elizabeth	Prisoner	F	16	1855	Kent	VIEW
WEEKS, Emma	Prisoner	F	28	1843	Surrey	VIEW
BENNETT, Elizabeth	Prisoner	F	19	1852	Surrey	VIEW
LEWIS, Elizabeth	Prisoner	F	19	1852	Norfolk	VIEW
HOUGH, James	Prisoner	M	21	1850	Surrey	VIEW
BAILEY, William	Prisoner	M	37	1834	Middlesex	VIEW
SCOTT, Samuel	Prisoner	M	21	1850	London	VIEW
SIMMONS, Charles	Prisoner	M	30	1841	Hampshire	VIEW
CLARK, William	Prisoner	M	28	1843	Lancashire	VIEW
GRIFFIN, John	Prisoner	M	52	1819	Surrey	VIEW
EAGER, Thomas	Prisoner	M	21	1850	Sussex	VIEW

Figure 4 (Colour online) 1871 Census, Ancestry (http://home.ancestry.co.uk/, 28 Apr. 2015).

In an average week, the Old Bailey Online attracts around 15,000 visits from dozens of countries around the world, and the vast majority are from private individuals accessing the site from outside of higher education. The first academic URL to appear in our lists normally comes in around thirty places from the top.[5] And it is not just usage of the internet that has changed historical practice. The four million plus people who watched episodes of the television series *Garrow's Law* over three seasons, the six million who tune in to each episode of *Who Do You Think You Are?* and the three million who sought out the first screenings of *Secrets from the Workhouse* are evidence of a vastly expanded audience for history, both online and

[5] See Sharon Howard, Tim Hitchcock and Robert Shoemaker, 'Crime in the Community Impact Report' (2010) (www.webarchive.org.uk/wayback/archive/20140614185215/http://www.jisc.ac.uk/media/documents/programmes/digitisation/analysis_cic.pdf).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press



Figure 5 (Colour online) *Charles Booth Online Archive*: Booth poverty map and modern map (http://booth.lse.ac.uk/cgi-bin/do.pl?sub=view_booth_and_barth&args=531000,180400,6,large,5, 28 Apr. 2015).

on TV.[6] And many of these 'viewers' are not just passively consuming history. They are doing their own research and writing.

Perhaps the best example of the involvement of a wider public in research and, via crowdsourcing, in the creation of new historical materials, is *Trove*. A resource which may be unfamiliar to many academic historians, *Trove* is the most successful historical crowdsourcing project in the world. It gives direct public access, and input, to around 400 million items held in the libraries of Australia, most notably the newspapers published up until the 1950s. On an average day, users of Trove make around 100,000 corrections to the newspaper archive – combining good citizenship with research of their own. In the process, *Trove* has built a community of historically interested users that significantly contributes to Australian public culture.[7]

Less dramatically, the same could be said of the online resources in the UK created by The National Archives and British Library, and to a lesser extent *Connected Histories*. The British Library's 'Million Images'

[6] For viewing figures, see 'Garrow's Law', *Wikipedia* (http://en.wikipedia.org/wiki/Garrow's_Law, 1 May 2015); 'Who Do You Think You Are? (UK TV Series), *Wikipedia* (http://en.wikipedia.org/wiki/Who_Do_You_Think_You_Are%3F_(UK_TV_series, 1 May 2015); and '"Secrets from the Workhouse" opens to over 3 million on ITV', *Digital Spy* (www.digitalspy.co.uk/tv/news/a493302/pbqJcUMHG6Nnr3, 1 May 2015).

[7] National Library of Australia, *Trove* (http://trove.nla.gov.au/, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

project has attracted a global community of active users.[8] A new audience of consumers and producers of history has evolved – many of them co-creating the sources of historical research, in the process of undertaking their own investigations.

The digital revolution has been a fantastic development for history, but it is not without its problems. The new resources threaten to deracinate the leavings of the past, and allow them to be used with little sense of context or meaning – all flashy quotes, located using keyword searches. But in our estimation, there has not been such a vibrant audience for history writing since the heydays of Macaulay and Gibbon.

Academic historians are obvious beneficiaries of this sea of change, but, with some exceptions, they have been rather reluctant to embrace and work with this new public of practising historians. Even when they have been willing, they have discovered that digital public engagement demands skills they do not have and resources they did not plan for. As with *Trove*, crowdsourcing has been used by a few academic projects. But in contrast to *Trove* – a project run by the National Library of Australia, an institution with public engagement embedded in its mission – academic projects have had some difficulty recruiting the help of a wider audience.

One of the most successful examples to date is *Transcribe Bentham*, which has invited public volunteers to transcribe the voluminous and often difficult to read papers of the utilitarian philosopher Jeremy Bentham, with the specific aim of breaking down traditional barriers between the public and academic research.[9] In five years, some 13,000 manuscripts have been transcribed or partially transcribed by volunteers, or 29 per cent of the 45,000 manuscripts which remained untranscribed at the start of the project. This is a tremendous achievement, but it has been a long and difficult journey. At this pace, the project will need another twelve years or so to complete the task. And although many people have participated, most have only worked on a few pages, while the vast majority of the pages have been transcribed by a small number of volunteers. The 'crowd' turns out to be rather small, and, given the fact all the transcriptions still needed editorial work, it is questionable whether this methodology has actually saved the project any money.[10] Indeed, saving money – as opposed to

[8] The British Library on flickr (https://www.flickr.com/photos/britishlibrary, 28 Apr. 2015). See also 'British Library uploads more than a million public domain images to Flickr', *Wired.co.uk*, 15 Dec. 2013 (http://www.wired.co.uk/news/archive/2013–12/15/british-library, 28 Apr. 2015).

[9] *Transcribe Bentham: A Participatory Initiative* (http://blogs.ucl.ac.uk/transcribe-bentham/, 28 Apr. 2015).

[10] T. Causer, J. Tonra and V. Wallace, 'Transcription Maximized; Expense Minimized? Crowdsourcing and Editing The Collected Works of Jeremy Bentham', *Literary and Linguistic Computing*, 27.2 (2012) (e-journal).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

deepening public engagement – may be the wrong motivation for using this methodology.[11]

Our own experience with crowdsourcing has been even less successful – when we invited people to contribute content to the *Old Bailey Online* through specially created wiki pages, the response was underwhelming, and eventually the wiki was discontinued.[12] We have had more success with very specific tasks, which take advantage of the way users are already interacting with the resource. When we added a simple correction feature to the Old Bailey site users responded with a variety of helpful comments on a wider range of topics than we expected. A similar approach was used with our parallel project, *London Lives*. The site provides access to over 240,000 manuscripts about poverty and crime in eighteenth-century London, and includes over three million separate name instances. We enabled registered users to link records together which they think concern the same individual, and so far some 3,000 of these 'sets' have been created.

The point is that crowdsourcing and public engagement are difficult, and require considerable skill, time and effort.[13] Academics cannot just assume that the 'public' will do what we ask them to do. In part, people are making their own histories, and do not necessarily want to be led by the academy. As both *London Lives* and the *Transcribe Bentham* projects have discovered, there needs to be a substantive dialogue between project staff and volunteers. A community of people working on the project needs to be created – so that both feel that they are getting something out of it, and the resulting resource is truly co-created. Unfortunately, current structures for funding projects, and allocating academic workloads, do not normally make the level of human resource necessary for this work available.

It is not just that academics have had limited success in creating a dialogue with family, local and public historians; there are also moves from within academic digital history which have had the effect of driving a wedge between historians working in higher education and a wider public. One of the virtues of digital history is that it has attracted scholars from different disciplines, with different skills, to look afresh at the evidence now that it has been transformed from mere words and objects into data. But the results of these interdisciplinary collaborations can be problematic.

[11] Trevor Owens, 'Crowdsourcing Cultural Heritage: The Objectives Are Upside Down' (2012), www.trevorowens.org/2012/03/crowdsourcing-cultural-heritage-the-objectives-are-upside-down/.

[12] Howard, Hitchcock, and Shoemaker, 'Crime in the Community Impact Report'.

[13] Sharon Howard, 'Bloody Code: Reflecting on a Decade of the Old Bailey Online and the Digital Futures of our Criminal Past', *Law, Crime and History*, 5.1 (2015), 21–4 (www.pbs.plymouth.ac.uk/solon/journal/vol5%20issue1%202015/Howard%20Bloody%20Code.pdf, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

Perhaps the best-known example is the Culturomics movement, emerging from the Cultural Observatory at Harvard, using Google's Ngram viewer to analyse word frequencies in Google Books. The Ngram viewer allows users to chart the relative frequency of words and phrases, year by year, in the full body of Google Books from 1800 to 2000. At its best, this forms a powerful way of exploring the content of what is now some 14 million volumes digitised by Google. But while the Ngram viewer is available to all, and easy to use and comprehend, some of the academic history that is being written on the basis of this tool can seem both divorced from context and increasingly arcane. The viewer has been used by its creators, Jean-Baptiste Michel and Erez Lieberman Aiden, to generate what they consider to be a newly 'scientific' reading of the past, that privileges varieties of history that are highly technocratic. Their most powerful claim to date is that the Ngram viewer demonstrates that irregular verbs in English have declined steadily over the last 400 years.[14] This is a potentially significant (if contentious) finding that implies language change is not subject to human agency. But whatever it is, it is not the stuff of popular history. Not only does the evidence produced ignore the limitations of the source collection (Google Books), but the mathematical methodology used and its results are too complex to be understood by most academic historians, let alone the wider public.

At its best, this sort of work can be deeply illuminating. Ben Schmidt's 'Prochronism' projects, for example, take the individual words in modern cinema and television scripts that purport to represent past events and compare them to every word published in the year they are meant to represent. In the process, he illustrates all the anachronism in *Downton Abbey*, and, more impressively, the subtle changes in the presentation of masculinity, decade by decade, in the evolving world of *Mad Men*.[15]

But this type of history moves the focus resolutely away from people like Sarah Durrant, and towards a variety of cliometrics – a 'scientific' approach to history that has little relevance for the new audience for history evidenced in popular culture. We are guilty of this too. One of our own projects, *Data Mining with Criminal Intent*, took us in precisely this direction.[16] In collaboration with Bill Turkel, at Western University (Ontario), we started to treat the Old Bailey text not as a collection of individual dramas, but as a 'massive text object'. Figure 6, for instance,

[14] Jean-Baptiste Michel *et al.*, 'Quantitative Analysis of Culture Using Millions of Digitized Books', *Science*, 331 (14 Jan. 2011), 177–8.

[15] Ben Schmidt, 'Prochronisms or: Downton Crabbey. Using TV anachronisms to learn about change in language' (http://www.prochronism.com/, 28 Apr. 2015).

[16] *With Criminal Intent* (http://criminalintent.org/, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

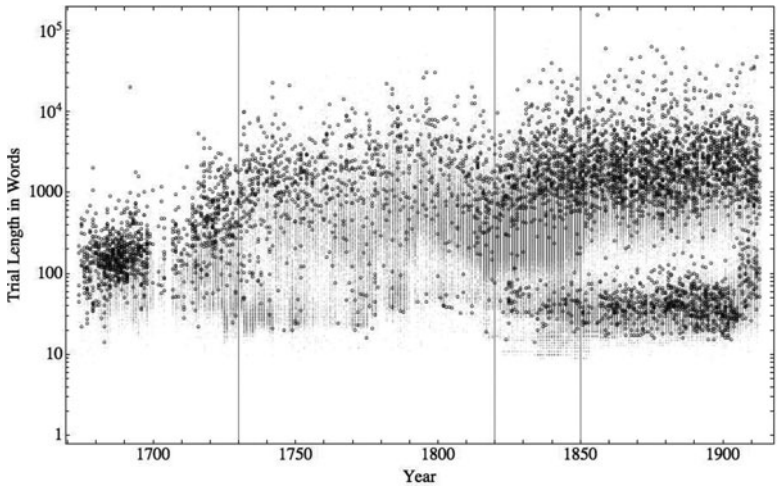


Figure 6 Distribution of trial lengths in words. Cases categorised as 'killing' displayed as black circles; all other trials as grey dots. 'Killing' includes all trials tagged for the offences of 'infanticide', 'murder', 'petty treason', 'manslaughter' and 'killing: other', by the Old Bailey Online (http://criminalintent.org/, 28 Apr. 2015).

represents all 200,000 trials in the Old Bailey Online, divided between those for forms of 'killing' and all other offences, and distributed according to how many words each trial contained – from the shortest, at 8 words, to the longest, at 157,000 words. By doing this, we discovered that the nineteenth-century trial came to be marked by large numbers of very short trials, as a result of the rise of plea bargaining, even for those accused of serious crimes like killing. For many, 'justice' had moved from the courtroom to the police cell.[17]

This graph provides a good example of how 'distant reading' allows us to discover new facts about familiar sources, but it did not result in popular history. And the same could be said of many of the important projects that are beginning to use sophisticated techniques such as formal network analysis, topic modelling, Text Frequency/Inverse Document Frequency measures and, most influentially, approaches derived from Bayesian probability – all generally thought of as forms of 'big data'

[17] 'Digging into Data Challenge Conference', *With Criminal Intent* (http://criminalintent.org/2011/06/digging-into-data-challenge-conference/, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

analysis.[18] The challenge is to link the individual to the complex patterns of data we can generate – to put Sarah Durrant back into the picture.

While some forms of academic *research* are in danger of turning their back on the rich opportunities digital history presents for bridging academic and popular history, academic history *writing* has been slow even to embrace the new possibilities of the digital as a means of dissemination. We continue to prioritise our traditional publication genres – the monograph, the journal article, the chapter in an edited collection. These are what we submit to the Research Excellence Framework (REF) in the UK, where in the most recent exercise websites and electronic databases accounted for just 0.5 per cent of the 6,431 submitted outputs in history.[19] They are also what we list on our CVs, and focus on when making decisions about promotion. We even resolutely ignore the electronic editions we consult when we write our footnotes and compile our bibliographies, choosing instead to cite the print versions. Similarly, we routinely fail to acknowledge our use of keyword searching as a research method.[20]

The 'e-book' is a good example of how slowly the academic history world, and its partners in the publishing sector, is changing. While academic monographs are now frequently published in e-book editions, alongside the traditional hardback and sometimes paperback, the e-book is typically little more than an online pdf document – taking advantage of only the most basic of the numerous opportunities the internet offers for making knowledge accessible and fostering dialogue. This failure to exploit the potential of the digital is particularly frustrating in the case of published editions of primary sources, such as tax and legal records and diaries. Despite the fact that these editions are often unrewarding to read in book form, and require very detailed indexes to be useful, they are still rarely delivered in an electronic format which would enable the keyword and structured searching necessary to maximise their usefulness.

[18] For a recent discussion of these approaches, see Tim Hitchcock, 'Big Data for Dead People: Digital Readings and the Conundrums of Postitivism', *Historyonics* (http://historyonics.blogspot.co.uk/2013/12/big-data-for-dead-people-digital.html, 28 Apr. 2015).

[19] Research Excellence Framework 2014: Overview Report by Main Panel D and Sub-panels 27 to 36 (Jan. 2015), pp. 51–2 (www.ref.ac.uk/media/ref/content/expanel/member/Main%20Panel%20D%20overview%20report.pdf, 1 May 2015).

[20] Jonathan Blaney, 'The Problem of Citation in the Digital Humanities', in Clare Mills, Michael Pidd and Esther Ward, *Proceedings of the Digital Humanities Congress 2012* (Sheffield: HRI Online Publications, Studies in the Digital Humanities, 2014) (www.hrionline.ac.uk/openbook/chapter/dhc2012-blaney, 17 July 2015); Tim Hitchcock, 'Confronting the Digital, or How Academic History Writing Lost the Plot', *Cultural and Social History*, 10.1 (2013), 12, 19.

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

In a book published in 2015, we tried to push the boundaries of the e-book, by designing it so that it is most productively read online.[21] Thousands of embedded hyperlinks take the user from relevant places in the text directly to the free online editions of the primary sources we have cited or quoted from, or the complete databases which underlie our tables and graphs, or catalogue entries of the printed primary source texts we cite, or e-book editions in Google Books of the secondary sources. These links are intended not only to make real the traditional purpose of a footnote – to allow the relevance of a specific piece of evidence to be confirmed by the reader, and the research journey of the authors to be made explicit – but also to encourage a new kind of reader engagement with our research. This is not a book that needs to be read sequentially from pages 1 to 450 (how many of us do that anyway?); instead, readers are encouraged not only to dip in and out of the book as they wish, but also to follow research threads back into the sources, where they can conduct their own research and consider other interpretations and lines of argument that we have not even thought of.

Bringing our publisher along on this journey proved difficult. While they were keen on the approach when we initially approached them, the production processes publishers use to generate books are not readily adaptable to new formats. They are still framed in terms of the production of a printed book, and the generation of basic pdfs from this printed text. It is difficult to work outside this system; the structures of academic publishing change very slowly.

While the online and the digital have fundamentally reshaped the landscape of historical research, bringing into being a new audience, and a new class of practitioners, the practices of academic research and writing threaten to ignore these new possibilities, heading in entirely different directions. Developments in the analysis of big data have tended to make digital history inaccessible, technocratic and in some ways irrelevant for a wider audience, while academic publishing is resistant to innovation. There are very real challenges to overcome if we are to achieve the potential for the internet to bridge the divide between academic and public history, and let history serve its primary function as a form of social memory, while also taking advantage of new methodologies created by big data. And in our estimation, the answer to this conundrum is provided by the internet itself.

First and most obviously, ease of publication means that historians can put the results of our research into the public domain almost instantly, and in innovative formats which are more readable by a wide audience.

[21] Tim Hitchcock and Robert Shoemaker, *London Lives: Poverty, Crime and the Making of a Modern City, 1690–1800* (Cambridge, 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

The academic blog is a good example of this. While some blogs explicitly address a public rather than academic audience (the 'History Matters' blog at the University of Sheffield is a good example of this), other historians use blogs to disseminate findings, or try out emerging arguments, for their peers as well as a wider audience.[22] Many, particularly younger, scholars are beginning to use social media and blogs as part of the process of developing ideas, collecting evidence and perhaps most importantly, ensuring that once complete, the history they have written has an audience of eager readers who have followed the research process from day one. One way of viewing the research blog is thus as the first draft of history; one can use one's blogs, and any comments received, as a starting point for writing more formal publications. This is the approach we are adopting on our latest project, the Digital Panopticon.[23]

Perhaps the best example of this is Ben Schmidt's hugely influential blog, *Sapping Attention*.[24] Schmidt's blog posts analysing nineteenth-century word frequency and authorship contributed to his doctorate, and will form part of his first book. Helen Rogers maintains two blogs: *Conviction: Stories from a Nineteenth-Century Prison*, on her own research; and a collaborative blog, *Writing Lives*, created as an outlet for the work of her undergraduates. These blogs bring together research and teaching, and in the process are building a substantial community of interest.[25] The list could go on. The *Many Headed Monster*, the collective blog authored by Brodie Waddell, Mark Hailwood, Laura Sangha and Jonathan Willis, is rapidly emerging as one of the sites where seventeenth-century British history is being re-written.[26] Others have a longer pedigree: Sharon Howard has been overseeing the *History Carnival* for over a decade, allowing historians from around the world to bring together themed content about early modern history.[27] For many historians, Twitter is where a lot of the ongoing discussion, in at least some corners of history, is taking place.[28] Following the hashtag #Twitterstorians, created by Katrina Gulliver,

[22] *History Matters: History Brought Alive by the University of Sheffield* (www.historymatters.group.shef.ac.uk/, 28 Apr. 2015).

[23] www.digitalpanopticon.org (1 May 2015).

[24] Ben Schmidt, *Sapping Attention: Digital Humanities: Using Tools from the 1990s to Answer Questions from the 1960s about 19th Century America* (http://sappingattention.blogspot.co.uk/, 28 Apr. 2015).

[25] Helen Rogers, *Conviction: Stories from a Nineteenth-Century Prison* (http://convictionblog.com/, 28 Apr. 2015); *Writing Lives* (http://www.writinglives.org/, 28 Apr. 2015).

[26] Broddie Waddell *et al.*, *The Many-Headed Monster: The History of 'the Unruly Sort of Clowns' and Other Early Modern Peculiarities* (https://manyheadedmonster.wordpress.com/, 28 Apr. 2015).

[27] Sharon Howard, *The History Carnival* (http://historycarnival.org/, 28 Apr. 2015).

[28] Laura O'Brien, 'Twitter, Academia and Me', *The Society for the Study of French History* (http://frenchhistorysociety.co.uk/blog/?p=348, 28 Apr. 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

which aggregates the 'tweets' of hundreds of individual historians, for a few days reveals a wild world of debate and engagement.

The relationship between these 'publications' and more formal academic outputs and reward procedures is unclear. At the moment, blogs listed on an academic CV are unlikely to be taken seriously as research publications, though they may well be seen as evidence of public engagement, which *is* starting to be recognised within promotions procedures.

Internet publication leads us to the controversial topic of Open Access, and the emerging requirements in the UK to make research publications freely available online. We do not propose to rehearse the various objections to Open Access, except to observe that it is hard to resist the conclusion that many academics have lost the plot on this issue. By focusing on economic costs, business models and copyright licensing, and insisting on ring-fencing the forms of dissemination used for traditional methods of scholarship, we are in danger of undervaluing the opportunities Open Access creates to widen the audiences for academic writing, to rethink the content of that writing and to develop new styles and genres – not just blogs, but more interactive or iterative forms of writing, where there is a dialogue between the writer and a broad audience.[29]

Many experiments in this area have failed to gain real traction, but journals such as *Digital Humanities Now* (http://digitalhumanitiesnow.org/) and the online platform *Open Library of Humanities* (www.openlibhums.org/about/) are pioneering new forms of cooperative peer review, pre-review publication and Open Access that point the way towards more useful, re-usable, transparent and open form of scholarly communication.

Innovative forms of communication can also be used to address the problems of the growing sophistication and lack of transparency of computer-based forms of analysis and the persistent divide between statistically and computer literate academics (still a small subset of the discipline) and everyone else interested in history. New presentational methods can be used to summarise both large bodies of data and the complex relationships identified through data analysis, and present them in an accessible form. A straightforward example is any graph produced by the Ngram viewer, which provides a visual summary of the language contained in millions of books. Even more user-friendly possibilities can be found via the in-vogue practice of 'visualisations' – pick up any newspaper today and you will find plenty of examples, presented with varying

[29] For a wider discussion of the role of social media in modern academic discourse, see Tim Hitchcock, 'Doing it in Public: Impact, Blogging, Social Media and the Academy', *Historyonics* (http://historyonics.blogspot.co.uk/2014/07/doing-it-in-public-impact-blogging.html, 1 May 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press



Figure 7 (Colour online) The distribution of punishment sentences for Old Bailey convicts, 1674–1709 compared to 1850–1913 (Digital Panopticon project: http://www.digitalpanopticon.org/datavis/punishment_treemap.html, 17 July 2015).

degrees of effectiveness. Used properly, visualisations (more precisely, infographics) provide an accessible form of 'distant reading', allowing anyone quickly to see broad patterns, identify trends and note anomalous cases – all of which can then be investigated in greater detail using a variety of research methods, including more traditional forms of 'close reading'.

For example, in the Digital Panopticon project, we are using visualisations to summarise judicial and penal experiences of felons in order to identify patterns in the thousands of life stories we are tracing of those convicted at the Old Bailey between 1780 and 1875. The project is still in its early phases, but Figure 7 illustrates, in the form of a 'treemap', an interactive visualisation of punishment sentences at the Old Bailey over the entire period of the *Proceedings*, demonstrating more effectively than a table or graph ever could the dramatically changing penal policy of Old Bailey judges. By using the sliding bar at the bottom of the page, the online user can watch the composition of punishment sentences change over the period from 1674 to 1913, revealing the transformation from

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

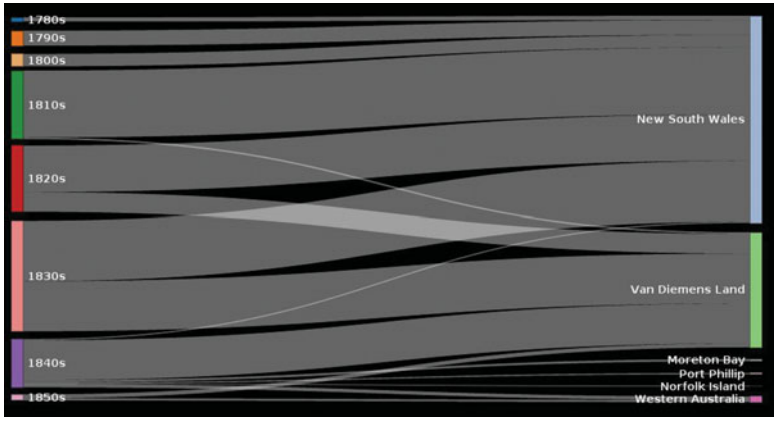


Figure 8 (Colour online) Destinations of transportees to Australia by decade. Each line represents one or more convicts, with the thickness of the line denoting the number of convicts in each decade who were sent to each colony (Digital Panopticon project: http://www.digitalpanopticon.org/datavis/sankey.html?data=linksdecade, 17 July 2015).

a penal system dependent on hanging, branding and whipping to one predominantly reliant on imprisonment.

The interactive user can then rapidly move from this 'distant reading' to individual stories – by clicking through from the relevant box to individual trials, to find Sarah Durrant.[30]

Another useful form of visualisation is a 'Sankey diagram', which maps relationships between two different variables. Figure 8 documents a key stage in the convict story – from sentence at the Old Bailey to transportation to Australia – showing how the date of conviction largely, but not exclusively, determined the colony to which convicts were sent. Ultimately, we will be charting the entire lives of these convicts (or as much as we can document), and our visualisations will chart various life courses from birth, through previous convictions, the Old Bailey trial and sentence, actual punishments experienced and subsequent life events (reoffending, marriage, death). When you are tracing the lives of some 90,000 convicts, the only way of summarising and analysing this evidence is through visualisations like this; and they should also be effective in communicating our findings to a wider audience.

[30]This feature is already available when using the statistical function on the Old Bailey Online: http://www.oldbaileyonline.org/forms/formStats.jsp (17 July 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

Mapping, using Geographical Information Systems (GIS), is another increasingly popular form of visualisation.[31] This kind of display has wide contemporary resonance as a result of the use of features such as Google Maps in day to day life, but it is also a widely used and accessible research tool. There are numerous examples, but to highlight a single one, the French Book Trade in Enlightenment Europe project spent some ten years developing interactive online maps that allow a user to interrogate the detailed records of a Genevan publisher and bookseller, the Société Typographique de Neuchâtel, from 1769 to 1794.[32] The Société sold books to purchasers all over Europe, and this website's mapping function (one of a number of types of visualisation on the site) allows the distribution of particular titles and authors to be mapped – reflecting a total of 70,000 sales transactions. Users can create their own maps by selecting particular book titles, authors, subject matter, types of publication, languages and client's professions, and limit by time period, thus allowing them to define their own research questions.[33] Unfortunately, it is not possible to drill down from the resulting maps to the individual records which form the underlying data.

Two terms have recently been widely used in relation to the internet – 'affordances' and 'disruption'. The first comes from the world of design, and refers to the uses that any bit of technology – any object – can be put. In writing history online, we are confronted with a remarkable set of new 'affordances', in terms of genres of publication and communication, as well as a wide variety of methodological approaches, from big data and distant reading to the simple but magical power of keyword searching on the infinite archive. But as an ever growing list of industries has discovered, every new affordance brings in its train 'disruption'. New career paths, Open Access, MOOCs and the empowerment of people outside the academy to produce their own histories, cut to their own cloth, challenges the historical profession to re-think how WE do history online. It is not a challenge we can afford to ignore.

Our belief is that we are in a fantastic age of new and popular historical engagement, and while it is not being led by academic historians (nor should it be), we need to be actively involved, and make sure that we add

[31] Ian Gregory, 'Using Geographical Information Systems to Explore Space and Time in the Humanities', in *The Virtual Representation of the Past*, ed. M. Greengrass and L. Hughes (Aldershot, 2008), 35–46.

[32] *FBTEE: The French Book Trade in Enlightenment Europe* (http://fbtee.uws.edu.au/main/, 28 Apr. 2015).

[33] Simon Burrows and Mark Curran, 'The French Book Trade in Enlightenment Europe', *Journal of the Digital Humanities*, 1.3 (2012) (http://journalofdigitalhumanities.org/1–3/the-french-book-trade-in-enlightenment-europe-project-by-simon-burrows-and-mark-curran/, 17 July 2015).

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press

our tuppence to the pot. As academics, we should do our bit to ensure that academic history is remade more open, more democratically accessible and ever more able to do the business of allowing society to question itself, to question its values in light of its past, its politics and its inherited principles. Despite the 'disruptions', as long as we keep in mind these underlying purposes of history writing, we can't go far wrong.

https://doi.org/10.1017/S0080440115000031 Published online by Cambridge University Press