Web as History

Brügger, Niels, Schroeder, Professor Ralph

Published by University College London

Brügger, Niels and Professor Ralph Schroeder.  
Web as History: Using Web Archives to Understand the Past and the Present.  
University College London, 2017.  
Project MUSE. muse.jhu.edu/book/81951.

For additional information about this book
https://muse.jhu.edu/book/81951

Peter Webster

Introduction

To anticipate the judgment of later historians on the very recent past has its risks. However, it may be that the decade following the turn of the millennium will come to be seen as a period marked by a shift in the nature of public discussion of the place of religion in public life in the UK. World events, indeed, made it likely that this should be so. The 9/11 terrorist attacks on the USA in 2001, the British involvement in the US-led war in Iraq in 2003 and the terrorist bombings in London in 2005 prompted an outpouring of anxiety in the media and in public life concerning the place of Islam in British politics and society.

These were not, however, irruptions in an otherwise stable field of discourse, for the period saw other significant changes in the religious landscape. The 2008 Criminal Justice and Immigration Act finally ended the statutory protection of Christianity from blasphemy (Kearns, 2008). The year before, the Labour government had signalled its willingness to relinquish its control over senior appointments in the established Church of England, with the report on The Governance of England. Legislation in 1999 removed hereditary peers from the House of Lords which for many left a job half-finished, since the position in the House of the bishops of the Church of England was still in question (Dorey and Kelso, 2011: 171–216). Taken together, these trends suggested that the uniquely privileged position of Christianity, and the Church of England
in particular, was under discussion to a greater extent than had been the case for decades.

The period also saw the coming to prominence of what has sometimes been termed the ‘New Atheism’: a polemically vigorous form of radical atheism that goes beyond a mere argument for a secular settlement in public life to a frontal assault on all forms of theistic faith. Prominent in this was The God Delusion by Richard Dawkins, which appeared in 2006, followed in 2007 by God is not Great: How Religion Poisons Everything by Christopher Hitchens (Amarasingam, 2010: 1–4).

To point out the coincidence of these three trends is not to assert any causal link between them. Such a determination must wait until the passage of time gives a longer perspective, and access to sources that now remain closed. But coincide they certainly did, and this chapter sets out to examine the potential of the archived web as a class of source material in which to observe the coincidence. Although there is a well-developed literature in the field of online religion, scholars have been slow to begin to exploit the archived web, as opposed to the live web (Campbell, 2011: 232–50). Whilst important groundwork has been done on the nature of the news media in an online environment (e.g. Burns and Brügger, 2012), there is also still room for studies of particular events and themes as they play out between mainstream media channels and the rest of the web.

Fully to examine the changing shape of religious discourse on the open UK web in this period would of course require a much larger study than the space available here would allow. This chapter confines itself to the exploration of a set of interrelated details of the larger picture, and by doing so proposes an approach to just such a larger study. All of the aspects treated relate to the unique position in British religious life of the archbishop of Canterbury, the leader of both the established Church of England and the global Anglican Communion.

Methods and sources

The primary data used for the research on which this chapter is based is the JISC UK Web Domain Dataset, held by the British Library. Acquired with funding from the UK agency the Joint Information Systems Committee, it is an extraction of all the resources in the Internet Archive from the country code top-level domain (ccTLD) for the UK (.uk) for the period 1996–2010. Two notes as to its contents are necessary. First, users of web archives must always deal with the fact that content can and does appear, change and subsequently disappear on the live web.
without having been visited by a crawler. In the case of this dataset in particular, comprehensive documentation of the crawl profiling – of matters such as how lists of seed URLs were compiled, how frequently and how deeply sites were crawled, and polices on deduplication of identical resources – is not available.\textsuperscript{2} The fact that a resource does not appear in the data cannot be safely read as indicating that such a resource was not in fact on the live web at the time in question, and the lack of understanding of the crawl profile means that it is difficult to hypothesize as to which content is more or less likely to be missing. As such, research questions must be framed in such ways as to avoid needing to equate an absence in the data with an absence in fact and then draw conclusions from the latter: in other words, to avoid the so-called ‘argument from silence’ (UK Web Archive, 2015a).

There are also significant limits on the scope of the data as a source from which to generalize about the whole experience of the UK. The criteria by which content should be included or excluded from a national domain crawl are expressed in the various implementations of Non-Print Legal Deposit in different nations, with varying treatments of ownership, geographical location and language. Be that as it may, there would be general agreement among web archivists that the ccTLD alone cannot encompass the whole of a national web sphere. In the UK, many organizations including political parties, banks, train companies and churches have non .uk domain names. Efforts to understand the scale of national web content that lies outside ccTLDs are in their infancy. However, a recent investigation by the British Library found more than 2.5 million hosts that were physically hosted in the UK without having .uk domain names (UK Web Archive, 2015b). As a result, the study presented here is cautious in making generalizations about the national web sphere for the UK from the more limited .uk data available.

The full JISC UK Web Domain Dataset is not available for use by individual researchers as a dataset, since it is some 32 TB in size and thus is unmanageable for the majority of users. However, this study makes use of a prototype user interface to a full-text index of the data – known as SHINE – made publicly available by the British Library.\textsuperscript{3} The UK Web Archive team have also placed in the public domain the Host Link Graph, derived from the larger dataset, which summarizes links between individual hosts in each year. The data appears in the following format:

\begin{verbatim}
2001 | host1.co.uk | host2.co.uk | 27
\end{verbatim}
This states that the data contains 27 individual resources from host1.co.uk that were crawled in 2001 and which contain one or more links to a resource at host2.co.uk.4

This chapter makes no use of the total numbers of linking resources per host given in the Host Link Graph, since to interpret them properly would necessitate an understanding of the total number of resources present on a host. A single link on a very small host might be thought to have a different significance than ten linking resources on a very large host. The chapter also does not analyse the significance of individual links, which are not recorded in the data at hand in a usable way. Instead, it focuses solely on host-to-host relations as a proxy measure of attention paid by the individual or organization by whom the linking host is controlled. Richard Rogers among others has explored the meaning of the link in a web context (Rogers, 2013: 39–59). The quality of that attention may of course be positive, negative or neutral; links may equally well be intended to draw the reader’s attention to content which the author deplores as much as to content (s)he endorses. In thinking about the nature of ‘hyperlink diplomacy’ between organizations, Rogers has characterized links as ‘cordial, critical or aspirational’ (Rogers, 2013: 45). This kind of close qualitative analysis of the sentiment of linkage is a matter for a larger study. Here, the concern is simply with attention.

The archbishops of Canterbury

The archbishop of Canterbury occupies a place in British public life for which it is difficult to find precise parallels elsewhere. As well as being the figurehead of the worldwide Anglican Communion, he is also leader of the Church of England, which is formally established as a state church whilst the Anglican churches in Scotland, Wales and Northern Ireland are not. Despite this, his position as the man who places the crown on the head of each new monarch of the United Kingdom, and as leader of the bishops in the UK parliament, has historically led many to regard him as in some poorly defined way the representative of all Britain’s Christians. As such, successive archbishops have understood it to be part of the role to intervene in matters of public controversy, even if some were less disposed to do so than others (Hastings, 1991: 84–98; Webster, 2015: 115–31). In their turn, for decades the mainstream media have tended to treat the interventions of the archbishop in different terms to those of the leaders of the other Christian churches and of the other faiths.
Rowan Williams and the 2008 sharia law controversy

Rowan Williams was named as archbishop of Canterbury in 2002 as successor to George Carey, and enthroned in 2003. Translated to Canterbury from the position of archbishop of Wales, Williams arrived with a reputation as an intervener in national affairs from the left wing of the political spectrum, having been shaped by the socialism of his native Wales (Shortt, 2008: 82–5). The essays collected in the 2012 volume *Faith in the Public Square* are directly concerned with the interplay of faith and politics. There was, however, one particular episode for which Williams may well be remembered for longest, and which prompted a significant change in the way in which his role was reflected in the UK web.

On 7 February 2008, Williams visited the Royal Courts of Justice in London to deliver a lecture to an audience of legal professionals, although admission was also open to the public. Among the members of the public in the audience was the present author, who remembers only a complex but clear and scrupulously balanced argument concerning the interaction of the secular civil law and religious law (and Islamic sharia law in particular), particularly in the area of the law of marriage. Put simply, Williams argued that many people who looked to religious principles to settle certain kinds of disputes did not find any reflection of that fact in the law, which contributed to their perception of marginalization. If this desire was unavoidable, it would be better to accommodate it within the law, and thus to a degree to control it, than to have it operate at a local level without any kind of restraint. There were other circumstances in which the law allowed parties in a dispute to go to arbitration without troubling the courts. Such an arrangement ought to be possible in these cases (Shortt, 2008: 390–402).

Unfortunately for Williams, the majority of the public were not at the lecture itself, and heard instead an interview given in advance to BBC radio, in which the archbishop suggested that some kind of accommodation of sharia was unavoidable (Internet Archive, 2008a). Although Williams’ time as archbishop up to this point had been dominated in the minds of church insiders by controversy over the ordination of gay clergy, the sharia law dispute brought him to public attention in a new way. All sections of the news media engaged with the story, some with outrage, and others with calls for the archbishop to resign, or for the Church of England to be disestablished. Williams’ predecessor George Carey described the suggestion as ‘disastrous’ in the tabloid *News of the World* (Webster, 2008). For many, the very limited accommodation that Williams proposed was lost in lurid
visions of stoning for adulterous women and the amputation of the hands of shoplifters.

Criticism also centred on Williams’ alleged naivety about the media and the likely reaction to the story (Goddard, 2013: 234–8). As an episode in media history, it was part of a recurring theme: the portrayal of senior religious leaders as well-meaning men who were either unaware of, or careless of the reaction which their interventions would provoke (De-la-Noy, 1990: 184–5; Webster, 2015: 125–7). In Williams’ case, a media narrative had already been established of his supposed intellectualism and inability to express ideas in a concise and clear way. Some of the staff at Lambeth Palace privately regretted that they had not enough time beforehand to digest the speech and its likely implications (Shortt, 2008: 401). Even sympathetic commentators were caught in the contradictory position of both praising Williams’ courage in raising a complex and emotive issue, whilst regretting that its expression had not been more easily digested by the media (Guardian, 2008a, 2008b).

There was also at the time some awareness that that supposed naivety in media handling extended particularly to the web. The experienced religion journalist Paul Vallely observed that ‘diligent website watchers’ had noticed the rapid online reaction:

‘as this crudeness of response was transmitted, and magnified, with increasing volatility by this new communications technology. [However] it seems there were no diligent website watchers at Lambeth. Or if there were, and they pointed out […] how seriously awry things were going, [Williams] failed to hear the electronic alarm bells. He would be a fool if he made the same mistake next time. And there will be a next time, make no mistake. Welcome to the world of the new media.’ (as quoted by Shortt, 2008: 401)

Reading a press storm in the web archive

Using the available data, it is possible to observe just this online reaction through the traces it has left. Extracted from the Host Link Graph dataset were all the occurrences of archbishopofcanterbury.org (the archbishop’s own site). Results which were outward links from captures of the archbishop’s site itself were removed, as were duplicates, where the Internet Archive had captured content from the same host more than once in a single year. In cases where there were multiple hosts that were part of a larger domain, these were not deduplicated. Although it would
have been straightforward to do so in the case of the larger media organizations such as the Guardian, which has multiple hosts (society.guardian.co.uk, education.guardian.co.uk, etc.) it was difficult to do so reliably for all such cases without examining individual archived pages, which was not possible at this scale. In any case, these accounted for less than 5% of the total. In the analysis that follows, it is assumed that a host abc.co.uk held the same content as www.abc.co.uk. It is also assumed that the Internet Archive was no more likely to miss hosts that linked to the Canterbury site than ones that did not. That is to say, if there are gaps in what the Internet Archive found (and there certainly will have been such gaps), there is no reason to suppose that they systematically skew this particular analysis.

In addressing the question of the sharia law controversy, it is convenient that it occurred very near the beginning of a calendar year (7 February). Although the absolute numbers are relatively small (347 for 2008), it is possible to see a significant rise in the total number of unique hosts found linking to the Canterbury site in 2008. The total for 2008 represents an increase of 49% on the previous year; it is 42% higher than the mean average of the three previous years; and it is 24% higher than the previous peak in 2004. An examination of the size of the dataset itself suggests that this is not to be accounted for by trends within the whole Host Link Graph, since the total number of linked pairs for 2008 is in fact considerably lower than any of the previous three years. A distant reading of the link graph therefore suggests that more attention was being paid to the Canterbury domain in 2008 than previously.

Of the hosts found linking to the Canterbury domain in 2008, some 153 (44%) were appearing in the data for the first time. Whilst always bearing in mind the fact that those hosts may actually have linked to the Canterbury domain before 2008 but were not captured by the Internet Archive doing so, what does the patterning in this group of hosts suggest about the degree and kind of attention being paid to the archbishop that year?

Mostly absent from this subset of the data were those hosts which formed part of the infrastructure of the religious (and indeed secularist) web. Very few national organizations within the Church of England or the other churches (and the ecumenical apparatus that links them) are to be found linking first in 2008, since most had begun to link before this date. The same can be said for the main secularist campaigning organizations such as the National Secular Society, and also for the mainstream media organizations.
Instead, the data show signs of a widening in the kinds of sites involved. Among those hosts linking for the first time, there was an academic journal dealing with the nature of the mass media, as well as mainstream social affairs thinktanks such as Demos. Also in the list are sites from within the fields of public relations consultancy, family law and political organizations such as the British National Party (of which more below). The Army Rumour Service, an unofficial and widely used chat forum for the British Army, contained an (overwhelmingly negative) thread about the matter from 9 February, two days after the lecture (Internet Archive, 2008b). However, the most significant group of hosts referring for the first time in 2008 are from the blogosphere.

An inspection of the blogs shows that while some of those recorded were written by Christians, or had established secularist or anti-religious themes, the majority had no particular religious agenda but were rather outlets for the miscellaneous opinions of their contributors. This would suggest that many bloggers who had previously not been particularly engaged with religious affairs in 2008 became interested enough to link to the archbishop’s domain. Some of the posts concerning Williams occur in blogs ostensibly dedicated to unrelated matters, such as that of an IT firm serving small businesses, which broke off from discussing instant messenger viruses to declare ‘God bless Rowan Williams’ (Internet Archive, 2008d). Some bloggers were positive; one, writing on 9 February, thought that the reaction was born of ‘deep prejudice and bigotry’ and that over time Williams might well be viewed as a ‘precipitator of a turning point in cultural reconciliation’ (Internet Archive, 2008c). However, the majority of the sample voiced similar sentiments to the more hostile voices in the mainstream media.

Of course, many of the largest and most commonly used blog platforms are not part of the UK ccTLD. That the apparent increase of attention to Williams in blogs hosted within the UK ccTLD is matched in those outside is indicated by one particular blog aggregation site, britishblogs.co.uk. The site was first captured by the Internet Archive in February 2006, at which time it referred to 16 posts tagged with the term ‘religion’. Before 2008, none of the content from the site captured by the Internet Archive contained links to the Canterbury domain. In 2008, by contrast, there were some 1,597 resources that did so. Even allowing for the very considerable probable levels of duplication involved in the way that sites such as this are crawled (with the same content captured multiple times in views by different subject tags), this suggests a step-change in the way in which Williams was being represented in the UK blogosphere.
York and Canterbury

England is unusual among those countries with an Anglican history in having not one but two archbishops, ostensibly of equal rank. When viewed in domestic terms, the archbishop of York has precisely the same authority within the northern province as does Canterbury in the south. However, the common media habit has for decades been to attend to public statements from Canterbury rather more closely than to those emanating from York.

Whether cause or consequence of this public and media perception, archbishops of York in modern times have tended to be less politically active, although there are exceptions to the pattern, most notably William Temple in the 1930s and 1940s. However, the most recent holder of the office, John Sentamu, has also not fitted the model, after succeeding David Hope in 2005. To take one measure of political activism, Hope had rarely intervened in the House of Lords and was indeed somewhat uncomfortable in principle with the role of the bishops in the House (Marshall, 2004: 67–8). Sentamu, by contrast, intervened regularly in the House and elsewhere on controversial matters from the outset. The most dramatic of Sentamu’s gestures was to cut up his clerical collar on BBC television in protest at the regime of Robert Mugabe in Zimbabwe (Sentamu is Ugandan by birth) (Internet Archive, 2007).

Sentamu, like Williams, also had a web domain dedicated to his work as archbishop. This first appears in the Internet Archive in October 2006 as a directory on the domain of the diocese of York before moving to its own domain (archbishopofyork.org) in early 2008 (Internet Archive, 2006a). Since Sentamu was arguably at least as politically active as Rowan Williams in the period under discussion, what does an analysis of inbound links to the York domain between 2008 and 2010 reveal about the relative attention paid to the pronouncements coming from the two men?

In order to address this, the method already documented in relation to Canterbury was repeated, with an extraction of all link pairs from the JISC Host Link Graph involving archbishopofyork.org, and a visual inspection and classification of all inbound referring hosts, using the live web or the Internet Archive. A total of 78 individual hosts were found, all but one of which it was possible to identify.

When compared with the same data for Canterbury, this absolute number of unique referring hosts was considerably smaller than the several hundred referring to Canterbury. Among them there are seven hosts
from the mainstream media, representing five organizations, including the BBC and several of the broadsheet newspapers. Predictably, slightly fewer than half were Christian organizations: parts of the national structure such as other dioceses, Anglican organizations within the diocese of York, and local congregations from around England.

The contrast between this data and that for Canterbury is in the relatively low number of inbound hosts from outside the churches and the media. None of the main campaigning secularist and humanist organizations are to be found linking to the York domain, and there are also few personal blogs. An examination of those individual bloggers shows that the references to Sentamu are often incidental, and do not demonstrate any sustained attention to the archbishop as a public figure. Despite Sentamu’s interventions in controversial national issues, then, the link structure of the UK web confirms the older pattern: that those outside the churches were still more likely to pay attention to the archbishop of Canterbury than to his northern colleague.

Changing patterns of religious discourse: a case study – the British National Party

British political parties of the far right have for long had a relationship of polemical tension with the archbishops of Canterbury. This has been the case since at least the late 1960s when Michael Ramsey was heckled by members of the National Front on account of his work on behalf of recent immigrants from the Commonwealth (Webster, 2015: 127–9). A consistent component of neo-fascist rhetoric has been a claim that the Christianity of the native English was under threat from uncontrolled immigration, and that the churches, and the established Church of England in particular, had colluded in allowing the crisis to arise (Jackson, 2010: passim). Whilst the British National Party has never won a seat in a UK general election to Parliament, it has at times achieved some success in elections to local government. After a period of apparent decline, the party enjoyed a significant increase in popular support in the years following the terrorist attacks on the USA in 2001 (Thurlow, 1998: 268–72). This rise in popularity coincided with a shift in the polemical strategy of the party, away from a traditional preoccupation with British Jewry and towards the ‘Menace of Islam’, in order to align the party more with public opinion (Copsey and Macklin, 2011: 85–6). This section considers the evolution of the website of the British National Party and its engagement with the archbishop of Canterbury.
The BNP site first appears at its current domain (bnp.org.uk) in the Internet Archive in 2001. This revamped version of the site, launched in July of that year, was evidence of a marked professionalization of the party’s mode of operation, with improved graphic design, use of PDF to deliver documents born in print, and the use of audio and video content (Copsey, 2003: 227–8). Using full-text search results derived from the SHINE interface provided by the British Library, this section examines resources from the BNP domain containing the search term ‘archbishop’. Whilst this result set could include references to other Anglican primates in Scotland or Wales, or indeed to their Roman Catholic counterparts, a qualitative examination of the archived content itself suggests that the prime concern of the BNP was with the Anglican archbishops in England, and Rowan Williams in particular.

Between 2001 and 2007 there were examples of an older complaint, familiar in conservative religious rhetoric generally, about a modernizing church leaving its traditional adherents behind. One article, first captured in 2003, connected the trend with the controversial appointment of Jeffrey John, an openly gay man, to be bishop of Reading, only for the decision to be reversed by Williams after sharp dissent from within the Church of England (Internet Archive, 2003). There were also early examples of BNP rhetoric shifting towards Islam. Already in 2001 the party was reporting the implementation of sharia law in Africa as evidence of the danger of Islamism (Internet Archive, 2001). The site also published an attack by a party activist on the Christian–Muslim Forum, a consultative group set up early in 2006, the launch of which Williams had hosted in the company of Prime Minister Tony Blair (Internet Archive, 2006b; Shortt, 2008: 336–7). That Williams himself was already a marked man in BNP circles was evident in an article published just weeks before the sharia controversy on proposals from the Labour government to repeal the historic laws protecting Christianity from blasphemy. BNP members should not expect Williams to fight this ‘constitutional vandalism’, it argued, since ‘Archbishop Rowan Williams is simply a Marxist in a dog collar, a man more interested in gay rights, immigrants and inter faith dialogue than in standing up for the Church of England’ (Internet Archive, 2008f).

So although the volume of content referring to the archbishops was limited before 2008, the themes that were to break into public consciousness in that year were nonetheless already evident in the BNP domain. However, 2008 saw a very significant increase in the amount of content containing the string ‘archbishop’ appearing in the archive. Between 2001 and 2007 there were fewer than 20 unique resources
found in the BNP domain containing the string ‘archbishop’; in 2008, the cumulative total (including those first published before that year) was in excess of 100. A proportion of the increased number of occurrences of the search string are accounted for by links to the initial post fed to other pages on the site. However, the bulk of the increase can be accounted for as new primary content, or new user comments on existing content. Some others contain references to other archbishops, such as to the alleged collusion of the Roman Catholic archbishop of Tours in the ‘betrayal of Charles Martel’, as the foundations of a mosque were laid in Tours, the scene of Martel's battle against Islamic invaders in 732 (Internet Archive, 2008g). However, the bulk of the increase can be attributed to the controversy.

The Internet Archive happened to crawl the BNP domain on 9 February, two days after Williams’ lecture at the Royal Courts of Justice. The site had carried a news item about the speech, summarizing it without particular comment, save for posting a link to the archbishop's site for users ‘to let him know how you feel about his comments!’ Within 48 hours the post had received 167 comments, almost without exception hostile to Williams’ argument. Alongside the merely vituperative, many of the commenters latched onto the episode as yet more evidence of the desertion of the Church of England of its historic duty to defend Christian England, and painted lurid pictures of beheadings and the amputation of hands on the streets of London. A number had evidently emailed the archbishop, and posted their email text as a comment (Internet Archive, 2008e).

From this point onwards, reportage on the BNP site in relation either to social cohesion at home or the consequences of alleged Islamic domination abroad was repeatedly connected to Williams and the 2008 controversy. If the authors of the content did not make the connection explicit, those adding comments very often did. Several of those commenting on the initial report thought that Williams’ comments were a publicity coup for the party (Internet Archive, 2008e), and it would seem that the level of engagement with the story from users of the site persuaded the party leadership that the issue should be pursued. Seizing on a particular phrase in an open letter from Williams to Islamic scholars, made public in July, party leader Nick Griffin was filmed in front of Lambeth Palace, issuing a call to British Christians to resist their own leaders’ collusion with the ‘Islamification’ of Britain (Internet Archive, 2008h). Even though the party had been publicly denouncing both sharia law and the archbishop before 2008, the evidence of the archived web shows that the controversy over Williams’ lecture led to a significant
upswing in content relating to both matters, both editorial and from users. Such a qualitative study of the evolution of specific hosts in the web archive affords the historian a new way of observing the development of religious discourse over time.

**Conclusion**

It is a brave historian who attempts to interpret the very recent past, as opposed to merely documenting it. As with most aspects of very recent history, the full significance of Rowan Williams’ lecture about sharia law will only become clear as the passage of time grants the historian a sufficiently long perspective from which to view it. An exhaustive qualitative examination of both the published record, and memoirs and private papers that are as yet inaccessible (not least the papers of the archbishop himself, not due to be released until 2038) will be needed to place the episode in its fullest context. Without these, we cannot yet know how changes in patterns of communication that are observable in the archived web were motivated, or how opinions expressed online related to broader patterns of social and intellectual change.

However, even if it is difficult to explain changing patterns of religious discourse on the web, we may nonetheless document those changes. First, the sharia law episode prompted a step-change in the levels of attention paid to the domain of the archbishop of Canterbury, as evidenced by the incidence of inbound links, and also a broadening of the types of hosts that contained those links. Second, a comparison of the inbound links to the Canterbury domain to that of the archbishop of York suggests that the historic privilege given to the views of Canterbury over those of York was extended onto the web. Regardless of their actual status in relation to each other within the Church of England, the media and the public at large seemed only to pay attention to Canterbury. Finally, a qualitative examination of the site of the British National Party shows that at least one organization, with a very particular concern with the place of Islam in British life, certainly took new account of the person of the archbishop as a result of the 2008 controversy.

This chapter has also sought to use the episode as a means of demonstrating both the potential for historians to utilize the archived web to address older questions in a new way, and some of the particular issues of method that web archives present. At one level, the methodological complications presented here – understanding the
meaning of a link from one resource to another, say – are peculiar to the archived web and must be understood anew. As with all other born-digital sources, there is work to be done among historians in understanding these issues of method, and in acquiring the skills needed to handle data at scale. At the same time, it is part of the historian’s stock-in-trade to assess the provenance of a body of sources, its completeness and the contexts in which those sources were transmitted and received. The task at hand is in fact the application of older critical methods to a new kind of source: a challenge which historians have confronted and overcome before.

This chapter has also tried to show some of the potential available to historians, should they accept the challenge. In the study of public controversy, the archived web allows the detection of changing communication patterns at scale that would be impossible using a traditional qualitative method. It also enables the detection of attention being paid online in places where a scholar would not think to look. More generally, the chapter has attempted to outline an approach that combines quantitative readings of the links in web archives with qualitative examination of particular subsets of resources. When dealing with a new superabundance of historical sources, a combination of distant and close reading will be required to understand the archived web.

Acknowledgements

My thanks are due to Ian Milligan, the editors and the anonymous peer reviewer for their comments on this chapter.