From Chartist Newspaper to Digital Map of Grassroots Meetings, 1841–44: Documenting Workflows

Chartism was the largest mass movement for democracy in nineteenth-century Britain. It is best remembered for its extraordinary tactics: 'monster' meetings of thousands of people in squares and fields; the three national petitions of 1839, 1842, and 1848, which gathered tens of thousands of signatures; and extraordinary events such as the 'risings' of 1839 and the 'plug plots' and conventions of 1842. Recently historians have reinterpreted the significance of the more ordinary and everyday elements of the movement. Malcolm Chase, Tom Scriven and others have shown how a familiar and quotidian culture was essential in sustaining Chartism in between the periods of mass agitation.[1] Historians of protest now take a more rounded and wide-ranging approach to understanding what adherence to the movement entailed.

An integral part of the organization of Chartism as a grassroots movement was weekly local branch meetings. Usually these meetings were held in the back room of pubs, but also in chapels, working men's halls, and increasingly as Chartists raised the money to build them, their own halls.[2] These meetings gave working men and (albeit in separate groups) women the opportunity to put their democratic principles into practice in voting, speaking, serving on committees and educating themselves. Eager to defend their legality, and to spread the word, the locations of the meetings were advertised in separate columns in the Chartist press, most notably

---

1. Malcolm Chase, *Chartism: A New History* (Manchester: Manchester University Press, 2007); Tom Scriven,

'Humour, Satire and Sexuality in the Chartist Movement', *Historical Journal*, 57.1 (March 2014), 157–78.

2. Katrina Navickas, *Protest and the Politics of Space and Place, 1789–1848* (Manchester: Manchester University Press, 2015).

in the *Northern Star and Leeds General Advertiser* newspaper (hereafter *Northern Star*). The paper was founded in November 1837 as the project of Chartist agitator and former Irish MP, Feargus O'Connor and the Leeds printer Joshua Hobson. It was published in Leeds and distributed nationally, reaching a regular circulation of 80,000 copies a week in 1839.[3] Aimed at a respectable working-class readership, the *Northern Star* followed the usual early Victorian newspaper mix of local, national and international news, but with an additional emphasis on advertising Chartist activities. There has not been a systematic analysis of the weekly meetings reported in the newspaper. How many meetings were there? Who held them and where? What can we learn by examining the geographic patterns of all the meetings that were held?

Historians still primarily understand Chartists on the basis of close readings of surviving texts rather than geo-spatial or social scientific modes of research. Indeed, Gareth Stedman Jones's essay 'Rethinking Chartism' in his highly influential collection, *Languages of Class*, inspired what became known as the 'linguistic turn' among scholars of early nineteenth-century British popular politics in the 1980s and 1990s. Attention to the texts of speeches and other literature became paramount to understanding the motivations and evolution of the movement.[4] And although important sources such as the *Northern Star* are now digitized and available online, as with most digitized newspapers, historians in effect use them in the same 'analogue' ways as they previously used microfilm or the original paper copies: reading one page at a time. A sea-change in research methods is occurring in that keyword searching is now the norm when using digital

---

3. 'Northern Star', Nineteenth Century Serials Edition, Birkbeck, University of London, and the British Library, beta version (August 2008) <http://www.ncse.ac.uk/headnotes/nss.html> [accessed 20 September 2016].

4. Gareth Stedman Jones, 'Rethinking Chartism', in *Languages of Class: Studies in English Working Class History, 1832–1982,* by Gareth Stedman Jones (Cambridge: Cambridge University Press, 1982), pp. 90–178; for work on Chartist texts see Mike Sanders, *The Poetry of Chartism: Aesthetics, Politics, History* (Cambridge: Cambridge University Press, 2009); Ariane Schnepf, *Our Original Rights of the People: Representations of the Chartist Encyclopaedic Network and Political, Social and Cultural Change in Early Nineteenth Century Britain* (Bern: Peter Lang, 2006).

resources. While this is undoubtedly positive, there are pitfalls to this new research landscape. Poor quality Optical Character Recognition (OCR) frequently forms the basis of the searchable text. If trusted blindly, the results of such searches may be incomplete or at worst: misinterpreted. In short, few patterns emerge if records are looked at sequentially; keyword searching is in effect still sampling with limited results. Nevertheless, the digital nature of the transcriptions in the *Northern Star* database opens up new possibilities if we are aware of the potential of digital analyses of texts that have hitherto only been read using conventional, micro-analytical approaches. The digitization of these periodicals and the development of text-mining tools to extract large amounts of quantitative as well as qualitative data from them, facilitates macro-analytical approaches. This article explores some of those possibilities by highlighting an approach that co-opts rudimentary linguistics and historical geographical approaches and applies them in a digital environment for the purpose of enhancing historical understanding. It does so by highlighting the workflow used by Katrina Navickas's *Political Meetings Mapper* project undertaken with the British Library Digital Scholarship Department.[5] The project started by seeking digital copies of the *Northern Star* newspaper, and ended with an interactive map of Chartist meetings. This map made it possible to understand the geographical and temporal distribution of grassroots Chartist activity for the first time. The result is a macroscopic view, giving what Katy Börner calls an opportunity to 'observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend'.[6] This is not a challenge to close reading, but a complement at a different resolution.

Workflow is of course always important to historians, but it finds itself in the foreground more often in some sub-disciplines than others. Digital history frequently asks historians to be

---

5. Katrina Navickas, 'Political Meetings Mapper', *British Library Labs* (2015)

<http://labs.bl.uk/Political+Meetings+Mapper> [accessed 20 September 2016].

6. Katy Börner, 'Plug-and-Play Macroscopes', *Communications of the ACM*, 54.3 (March 2011), 60–69.

critical and indeed open about their sources and methods; however, digital history is not alone, nor did it invent the in-depth discussion of methodology and workflow. For example, E.A. Wrigley's *The Early English Censuses* (2011) is a book about the workflows the author used in his analyses of these early censuses, building upon decades of research in historical demography. The book provides such a clear map for readers of what the author did to the records, that one could call it a monograph on historical workflow.[7] Likewise, much of the work presented in journals such as *The Economic History Review* also focus on processing data through mathematical models that are meticulously described so as to be reproducible.[8] This social-scientific approach to reproducibility and transparency is an offshoot of the scientific method, which few humanities scholars have found a need to emulate until recently. This shift towards the scientific method may in part be explained by the fact that 'digital' analyses are often actually interdisciplinary uses of social scientific methods. Both mapping and linguistics are social scientific approaches to knowledge building which have recently become accessible to humanities scholars in the form of digital tools and through new publications such as *The Programming Historian* (2012–Present), as well as *Exploring Big Historical Data: The Historian's Macroscope* (2015), which have taken the lead on prioritizing reproducibility in humanities research.[9] This article builds on the work of *The Programming Historian* and reproducible research practices, generalizing the processes used by Navickas so that they can be useful to scholars working on different types of records but with similar aims of acquiring, cleaning, geocoding, and displaying historical information from across

---

7. E.A. Wrigley, *The Early English Censuses* (Oxford: Oxford University Press, 2011).

8. *The Economic History Review* (1927–Present).

9. Adam Crymble, Fred Gibbs, Allison Hegel, Caleb McDaniel, Ian Milligan, Evan Taparata and Jeri Wieringa, eds, *The Programming Historian*, 2nd ed. (2016) <http://programminghistorian.org/> [accessed 20 September 2016]; Shawn Graham, Ian Milligan and Scott Weingart, *Exploring Big Historical Data: The Historian's Macroscope* (London: Imperial College Press, 2015).

Author
Deleted: Katrina Navickas

4

a set of historical primary sources.[10]

Acquire

As yet, digitizing a large historical corpus is impractical for most individual historians. Even a publication run on the scale of a newspaper like the *Northern Star*, which was published for a modest fifteen years between 1837 and 1853, still requires a library partner in possession of the paper or microfilm copies, at the very least. For most scholars, acquiring a newspaper or similarly substantial digital corpus involves finding one that has already been digitized.

As Tim Hitchcock notes, much of that work has been done by private companies who charge subscription access to material.[11] The 2014 change to UK copyright legislation has gone a long way to facilitate greater access to digital corpora for UK-based researchers. The new law gave researchers the right to make copies of any textual records for which they had 'legal access', and made unenforceable any terms of use that prohibit the making of copies for non-commercial text and data mining analysis.[12] The result of this has been a new openness by many commercial publishers to provide limited access to certain researchers as they try out this new model of access to their records.[13] However, for most scholars – particularly early career scholars,

---

10. Katrina Navickas, *Political Meetings Mapper* (2015–2016) <http://politicalmeetingsmapper.co.uk> [accessed 27 May 2016].

11. Tim Hitchcock, 'Privatising the Digital Past', *Historyonics* (2 June 2016) <http://historyonics.blogspot.co.uk/2016/06/privatising-digital-past.html> [accessed 20 September 2016].

12. 'Exceptions to Copyright: Research', *Intellectual Property Office, UK,* (October 2014) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/375954/Research.pdf> [accessed 14 July 2016].

13. 'Gale Leads to Advance Academic Research by Offering Content for Data Mining and Textual Analysis', *Cengage Learning* (17 November 2014) <http://news.cengage.com/higher-education/gale-leads-to-advance-academic-research-by-offering-content-for-data-mining-and-textual-analysis/> [accessed 14 July 2016].

independent scholars, or postgraduate students – getting a positive response still involves a level of privilege that is important to recognize. In the case of the *Political Meetings Mapper* project, access to the textual layer of *Northern Star* database was granted by British Library Labs, whose mandate is to promote the use of digital resources in the library collection.[14]

Each request will be met differently by the owners of the data, and it is not uncommon to be asked to pay fees or negotiate legal nondisclosure agreements. It is also not uncommon for requests to be rejected outright or ignored. Sometimes these requests will be refused on technical grounds. What seems like a simple request for information may require someone to spend considerable time figuring out how to get what you want to use and package it in a way that makes it easy to transport. Even small collections, if poorly documented or without an individual on the team who knows how the system works, can be difficult to extract from their databases. Asking for data is an art rather than a science; however, as Christian Kreibich notes, there are strategies for improving one's chances of success, ranging from using a university email address to emphasize the professional nature of the request, to being clear about why one wants the data, and of course expressing one's gratitude.[15]

In the case of the *Political Meetings Mapper* project, Navickas submitted a proposal to and won the second *British Library Labs Competition*, which is funded by the *Andrew W. Mellon Foundation*, and which awards two scholars per year with privileged access to digital collections in the British Library collection as well as access to library expertise. As part of that competition, Navickas was given access to the collection via a series of digital indexes. With these, it was possible to identify the filenames of relevant page scans (Figure 1), and a set of the Extensible Markup Language (XML) files that contained the searchable text layer (Code Block 1), both of

Author
Deleted: Figure
Author
Deleted: 2

---

14. 'British Library Labs', *The British Library* <http://labs.bl.uk/> [accessed 20 September 2016].

15. Christian Kreibich, 'How to ask for datasets', *Medium.com* (30 April 2015)

<https://medium.com/@ckreibich/how-to-ask-for-datasets-d5ef791cb38c#.b02iufreo> [accessed 3 June 2016].

which had to be manually downloaded. This process took approximately sixteen hours of work. The dataset included over 1,700 page scans of 208 issues of *Northern Star* between 1841 and 1844, with a total word count of around 312,000 words for the column of interest: 'Forthcoming Meetings'. Navickas chose this sample date range due to the time constraints of the project and because it covered the most active period of the Chartist movement. The page scans for the most important year in Chartist history, 1842–3, were not available in the collection so these had to be accessed manually from Gale-Cengage's database, *Nineteenth Century Newspapers*.[16]

INSERT FIGURE 1 ABOUT HERE

Figure 1: Page scan extract from front page of The *Northern Star* newspaper, 9 February 1839, © British Library, WO1_NRSR_1839_02_09-0001.tif. Reproduced with permission of the British Library.

```
<typeOfPublication>Newspaper</typeOfPublication>
<subCollection>Regional Weekly</subCollection>
</title_metadata>
<issue_metadata>
<volumeNumber></volumeNumber>
<issueNumber>65</issueNumber>
<printedDate>SATURDAY, FEBRUARY 9, 1839</printedDate>
<normalisedDate>1839.02.1839</normalisedDate>
<pageCount>8<pageCount>
<reelID>0112</reelID>
<qualityRating>Fair</qualityRating>
</issue_metadata>
<pageImage>
<pageSequence>0001</pageSequence>
<pageImageFile>W01_NRSR_1839_02_09-0001.tif</pageImageFile>
<pageCoordinates>911,464,5342,7209</pageCoordinates>
<pageSkew>50</pageSkew>
</pageImage>
<pageText>
```

16. 'British Library Newspapers', *Gale Cengage* <http://gale.cengage.co.uk/british-library-newspapers.aspx> [accessed 20 September 2016].

```
<pageWord coord="94,1002,304,1047">TICTORIA</pageWord>
```

Code Block 1: XML extract of the text layer of *Northern Star* newspaper, 9 February 1839. Much of the XML refers to the pixel coordinates where the word can be found on the original page scan. This is used to highlight keywords when using the commercial provider's website.

Clean

In order to map the Chartist meetings, the next step involved identifying relevant articles in the newspaper. The project focused only on one column, the 'forthcoming meetings' column of the *Northern Star,* as this provided the most succinct and regular form of wording and punctuation that could be most efficiently extracted without having to sift manually through extra contextual narrative description. Initially this was identified through the standard column heading, 'forthcoming meetings', but as it quickly became clear that this was usually on the same page, Navickas began to isolate the column manually (Figure 2).

INSERT FIGURE 2 ABOUT HERE

Figure 2: Excerpt of 'Forthcoming Chartist Meetings', *Northern Star*, 13 February 1841, © British Library, WO_NRSR_1841_02_13-0005.tif. Reproduced with permission of the British Library.

Given the relatively modest size of the collection, a manual approach proved more effective than keyword searching. This particular newspaper had been digitized in the previous decade using the latest OCR software available to create the searchable text layer that is stored in the XML files. The problems of poor quality OCR are well documented by Holly Rose, who noted in 2009 that a sample of Australia's massive *Trove* newspaper database contained accuracy

levels ranging from 71 to 98 per cent, with 71 per cent accuracy representing 145 errors in an average paragraph of text.[17] These results were comparable to those found by the Koninklijke Bibliotheek in 2008.[18] The accuracy levels of *Northern Star* newspaper OCR-generated transcriptions are unknown, as quantifying this measure requires manually counting errors in a sample of pages. However, even when the relevant column had been identified, it quickly became clear that the text contained enough errors that it could not be relied upon for a systematic extraction of Chartist meetings within the column (see Code Block 2).

```
<pageWord coord="90,2376,232,2416">&amp;suraucee</pageWord>
<pageWord coord="208,2377,297,2414">may</pageWord>
<pageWord coord="285,2377,338,2414">be </pageWord>
<pageWord coord="326,2377,338,2414">effected,</pageWord>
<pageWord coord="440,2374,547,2411">Daily</pageWord>
<pageWord coord="93,2415,250,2458">&apost;opeetuses</pageWord>
<pageWord coord="236,2419,328,2456">may</pageWord>
<pageWord coord="317,2418,375,2455">be </pageWord>
<pageWord coord="362,2416,432,2455">had</pageWord>
```

Code Block 2: Example of XML errors on key terms that made it impractical to re-use the original XML.

Originally, the project envisaged setting up a crowd-sourced transcription site, building upon the model used by Bob Nicholson's *Victorian Meme Machine,* which would have required

---

17. Rose Holley, 'How Good Can It Get? Analysing and Improving OCR Accuracy in Large Scale Historic Newspaper Digitisation Programs', *D-Lib Magazine*, 15.3/4 (March/April 2009) <http://www.nla.gov.au/ndp/project_details/documents/ANDP_HowGoodCanitGet.pdf>

18. Edwin Klijn, 'The Current State-of-art in Newspaper Digitization', *D-Lib Magazine*, 14.1/2 (January/February 2008) <http://www.dlib.org/dlib/january08/klijn/01klijn.html> [accessed 20 September 2016].

volunteers to transcribe the columns by hand.[19] However, it proved more economical and efficient to perform the OCR again using the latest version of commercial OCR software.[20] This provided new transcriptions with approximately eight out of ten words transcribed correctly – a much greater level of accuracy than the text in the XML files. The results were then cleaned up by a small team of research assistants: Samantha Walkden, Megan Dibble and John Levin, who checked and corrected the OCR files. The corrections mainly involved altering spacing and punctuation of the columns. This amounted to twelve days of work and resulted in four years of newspaper transcriptions (1841–44), saved in .txt format.[21] The new OCR'd copy of the transcriptions was now suitable to be used for research.

```
1. London – The public Discussion will be resumed in the City Chartist
2. Hall, 1, Turnagain-lane, on Sunday next, at half-past ten o'clock in the
3. forenoon. At three o'clock in the afternoon of the same day, the
4. Metropolitan Delegate Council will assemble for the dispatch of
5. business. – In the evening at seven o'clock, Mr. J. H. R. Bairstow will
6. deliver a lecture.
```

Code Block 3: Text file of the OCR'd newspaper text, *Northern Star*, 23 November 1844.

Author
Deleted: 5

Extract

With digital text clean enough to identify relevant entries reliably, the next step was to extract those entries and structure them in a way that would make it possible to map the location of meetings. At this stage the need was to find any mention of a meeting and save the result to a database. There are a number of ways this could have been achieved. The *Political Meetings Mapper* project chose to use some custom gazetteers compiled by Navickas that contained words known

---

19. Bob Nicholson, 'Introducing... the Victorian Meme Machine', *Digital Victorianist* (18 June 2014)

<http://www.digitalvictorianist.com/2014/06/victorian-meme-machine-interviews/> [accessed 14 July 2016].

20. The project used *Abbyy FineReader 12*, a commercial OCR package.

21. 'Text File', *Wikipedia* <https://en.wikipedia.org/wiki/Text_file> [accessed 14 July 2016].

to frequent that weekly column of meetings. This gazetteer was a simple text file with one term (lower case) per line.

For historical projects there is an added challenge: many of the individual pubs, halls and some streets of the 1840s no longer exist. To solve this problem, Navickas identified locations manually, using historic trade directories digitized by the University of Leicester and looking visually for the sites on historic town plans.[22] The information from the trade directories was obtained from the images rather than from any underlying XML. Using old town plans, it was possible to georeference an old map and put it into Google Earth, where it could then be used to find the current geo-coordinates for those lost places.[23] Therefore, the research, like many small-scale digital projects, could not be done through a 'one-stop shop' software package, but involved the careful curation of various ready-made, custom-built, proprietary and open access resources. This process also raises questions about sustainability and replicability, as many of these resources rely on institutional hosting or, commercial tools such as Google Earth or Fusion Tables, require signing up to online accounts and uploading one's data to their servers.

The gazetteer was then used to search the text for matches. This was done using a custom Python programme by Adam Crymble, 'Using Gazetteers to Extract Sets of Keywords from Free-Flowing Texts', which is described as a step-by-step tutorial on *The Programming Historian*.[24] Navickas adapted this code slightly for the project's needs, but the core principles behind the original tutorial apply to the needs of the workflow herein described. The full code (hereafter 'the Python code') used by Navickas can be found on *Zenodo* in the Project's

Author
Deleted: Crymble

Author
Deleted: Crymble

22. 'Historical Directories of England and Wales', *Special Collections Online, University of Leicester* <http://specialcollections.le.ac.uk/cdm/landingpage/collection/p16445coll4> [accessed 14 July 2016].

23. *Google Earth*, 2001–Present <https://www.google.co.uk/intl/en_uk/earth/> [accessed 20 September 2016].

24. Adam Crymble, 'Using Gazetteers to Extract Sets of Keywords from Free-Flowing Texts', *The Programming Historian* (2015) <http://programminghistorian.org/lessons/extracting-keywords> [accessed 20 September 2016].

repository.[25] This script was run on each column of the meetings' announcements in turn, extracting the text relevant to a single meeting as it went.

Geocoding

At this point in the workflow, the individual meetings had been identified. The next step was to geocode the meeting locations. Geocoding is the process of pairing words that relate to a physical location, to coordinates that represent the same place on a map. There are a growing number of tools that can perform this task, however these change frequently as new software emerges, and so it is more important to understand what geocoding does to the historical data.[26] It is a process that involves converting strings of text that refer to places such as 'China Walk, Lambeth' to its decimal latitude and longitude (51.495397, -0.1126751). There are a number of formats for geocoding that go beyond latitude and longitude, and each system of mapping has strengths for a particular area. For example, the British National Grid is commonly used to study the geography of Britain as it provides a highly accurate representation of British places, but the further from the British archipelago one travels the more distorted the results.[27] This is in part caused by the challenge of rendering the curved surface of the globe onto a two-dimensional map. Readers are advised to consult with a subject specialist in geography or cartography on which geocoding format is most appropriate for their project. As this project intended to use

25. Katrina Navickas, Ben O'Steen and John Levin, 'meetingsparser: package', *Zenodo* (2016) <https://zenodo.org/record/57875#.V4eP9tAtL20> [accessed 20 September 2016], DOI: 10.5281/zenodo.57875, <https://github.com/BL-Labs/meetingsparser/tree/concise-version> [accessed 20 September 2016].
26. The town names were geo-coded using *IDRE Sandbox* <https://sandbox.idre.ucla.edu/sandbox/sandbox-geocoder> [accessed 20 September 2016], then the co-ordinate information for the historic addresses was added manually (process described above) to the gazetteer generated by the geocoder.
27. 'The National Grid', *Ordnance Survey* <https://www.ordnancesurvey.co.uk/resources/maps-and-geographic-resources/the-national-grid.html> [accessed 14 July 2016].

Omeka to display the data (see below), Navickas chose to use latitude and longitude because this format was required for the Omeka maps plug-in.[28] Geocoding was conducted by the Python code at the same time as the extraction process identified above, however, from the perspective of a workflow this is a separate step. The geo-coordinates were then manually saved to the CSV file beside each entry, with latitude and longitude each in its own column (Figure 3).

INSERT FIGURE 3 ABOUT HERE

Figure 3: CSV file containing each meeting and its associated metadata. Dublin Core is the metadata standard required for the Omeka content management system (<http://dublincore.org/> [accessed 20 September 2016]).

Dating the meetings

As each meeting also took place at a certain time, and the temporal distribution of meetings undoubtedly had historical meaning, it was important to identify the meeting date. As noted, each meeting had a place and a time listed in the advert in the *Northern Star* newspaper. Unfortunately, the dates were not written to be easily machine-readable. It was common, for example, for a meeting to be listed as 'this Thursday' or 'tomorrow'. Because the *Northern Star* was always published on a Saturday, and because we know the date each newspaper issue was printed, it was possible to convert phrases like 'tomorrow' into the date of the meeting referred to using some simple Python code that employed pattern matching using regular expressions.[29]

---

28. Anon, 'Geolocation plugin for Omeka', Version 2.0 <https://omeka.org/codex/Plugins/Geolocation_2.0> [accessed 14 July 2016].

29. For an introduction to Regular Expressions, see Doug Knox, 'Understanding Regular Expressions', *The Programming Historian* (2013) <http://programminghistorian.org/lessons/understanding-regular-expressions> [accessed 20 September 2016]; Laura Turner O'Hara, 'Cleaning OCR'd text with Regular Expressions', *The*

This list was manually created for the needs of the current project. Once dates had been identified, they were added to a new column in the CSV file described above. At this stage of the workflow, all information required to map the meetings over time had been extracted and structured.

Display

The final step was to import the geo-coded meetings in the project website's digital map. The project used Omeka and the 'Geolocation' plug-in. Omeka is a free content management system for building websites produced by the Roy Rosenzweig Center for History and New Media at George Mason University. It was originally designed for the gallery, library, archives, and museum industry as a means of producing exhibits of collections. It has strengths for those seeking to batch upload items that include metadata (such as museum objects). The project has a number of plug-ins that add functionality to the site, including mapping locations as used in this project. Omeka has some limitations from a user perspective, such as an inability to export search results of all meetings from a particular locale, for example. There are alternative websites and content management systems that could be used for similar projects, and the reader should consider the most suitable and sustainable platform for their project needs and audience.[30]

As Navickas planned to use this plug-in to build a digital map, the above steps were designed so that the data created would be compatible with this tool. This included adding a column, which specified the optimal zoom level of the map for display. Import was conducted using the instructions for the plug-in. The result was a digital map of 4,962 Chartist meetings

Author
Deleted: Crymble

between 1841 and 1844, which can be viewed on the project website.

To provide historical context to the landscape, Navickas overlaid a nineteenth-century map of Britain over the modern Google Map used by the plug-in. The most easily available large-scale map was the first-edition Ordnance Survey map of the UK (1885), through the National Library of Scotland's Application Programming Interface (API) service.[31] This API was compatible with the Geolocation plug-in through an intermediary service, 'Leaflet', which enabled the historic map to be tiled, layered, and displayed at different levels over the Google Map (see Figure 4).[32] Readers need to consider the sustainability of third-party programmes for display and visualization. In July 2016, *Mapquest*, the service providing the background map tiles was discontinued, resulting in the base map becoming unavailable (see Figure 5).[33]

INSERT FIGURE 4 ABOUT HERE

Figure 4: Political Meetings Mapper Geo-location plug-in map, using Geolocation plugin for Omeka (http://omeka.org/add-ons/plugins/geolocation/) and Leaflet JavaScript Library (http://leafletjs.com/) . Meetings locations plotted on first edition one-inch to the mile Ordnance Survey map of the United Kingdom, 1885-1900, using the National Library of Scotland API under a Creative Commons Attribution 3.0 Unported Licence

---

31. 'NLS Historic Maps API - Historical Maps of Great Britain for Use in Mashups', *National Library of Scotland* <http://maps.nls.uk/projects/api/> [accessed 14 July 2016].

32. 'Leaflette Javascript Library' <http://leafletjs.com/plugins.html>; the code for the amended plugin is available at <https://zenodo.org/badge/latestdoi/23273/BL-Labs/Geolocation>, DOI: 10.5281/zenodo.57877 [accessed 20 September 2016].

33. Lori Colston, 'Modernization of Mapquest results in changes to direct tile access', *MapQuest + Developer Blog* (15 June 2016) <http://devblog.mapquest.com/2016/06/15/modernization-of-mapquest-results-in-changes-to-open-tile-access/> [accessed 14 July 2016].

Author
Deleted: Crymble

Author
Deleted: 7

Author
Deleted: 8

Author
Deleted: 7

Author
Deleted: 7

Author
Deleted: from

(http://maps.nls.uk/projects/api/)

INSERT FIGURE 5 ABOUT HERE

Figure 5: Political Meetings Mapper map, with missing base map tiles caused by a change in the terms of use by Mapquest that unexpectedly affected the project, 11 July 2016. API at http://maps.nls.uk/projects/api/ and used under a Creative Commons Attribution 3.0 Unported Licence. This demonstrates a clear lesson in digital sustainability.

Conclusion

In this project we have learned the advantages of taking a digital approach to newspaper sources. To take one example from the Chartist meetings column of the 22 January 1842 issue of the *Northern Star*, the Red Lion public house in Golden Square, London, advertised its forthcoming meeting the following Saturday, a spirited lecture by Mr L.H. Leighs denouncing 'free trade fallacies'.[34] Using the digital project, historians can not only find out that Chartist groups were also gathering on that evening elsewhere in London in the Hit or Miss public house in Mile End and in the Black Bull, Hammersmith (to celebrate the birthday of Thomas Paine), as well as all around the country. But the project database also displays the much wider context for these meetings situated in place and time. Historians can discover different and much broader connections than they could do manually. How common were these meetings in those particular places? How were they spread across the city, and how did this change over time? Of course, the historian can answer some of these questions using traditional approaches. However, using digital methods enables them to support their conclusions with more confidence, with a sample

---

34. 'Red Lion, King-street, Golden-square', *Political Meetings Mapper* <http://politicalmeetingsmapper.co.uk/maps/items/show/22801> [accessed 27 May 2016].

of 5,000 meetings rather than say a hundred, and in a format that appeals to our visual and spatial faculties. So, for example, the data clearly displayed the wide distribution of Chartist meetings across London.

London Chartism has been curiously under-studied compared to other regions of England, with the last major study of the metropolitan movement being David Goodaway's *London Chartism, 1838–1848* (1982).[35] Mapping the meetings' data showed the spread of Chartist branches and meeting sites across the city (Figure 6), with particular concentrations in Soho, Shoreditch-Spitalfields and Southwark. It also demonstrated the concentration of trades' branches in particular areas. For example, the tailors had several Chartist branches in Soho and the West End, where their trade worked and lived (Figure 7). The map confirmed the impression of London as an artisanal and trades-based movement with easy access to familiar and close-by meeting sites related to their trades' activities (many of the sites were pubs also holding the box for their friendly societies and trade unions). Chartist activities could therefore be characterized as part of the everyday rather than the extraordinary, drawing their strength from locality and proximity as well as from a wider delegate system across the city.

INSERT FIGURE 6 ABOUT HERE

Figure 6: Heat-map of concentration of London meeting sites in *Northern Star*, 'Forthcoming Meetings', 1841–1844, created using QGIS.

INSERT FIGURE 7 ABOUT HERE

Figure 7: Chartist tailors' meeting sites plotted on extract of Richard Horwood's map of London,

---

35. David Goodaway, *London Chartism, 1838–1848* (Cambridge: Cambridge University Press, 1982).

1792, British Library, Maps.Crace.v,

<http://www.bl.uk/onlinegallery/onlineex/crace/p/007zzz000000005u00173000.html>, geo-referenced and layered on Google Earth [accessed 20 September 2016].

The project also gave an insight into the history of the newspaper and its reach in particular. Plotting the meeting advertisements showed that even though the *Northern Star* was published in Leeds, the spatial distribution of reporting in the paper was not just concentrated in the West Riding of Yorkshire and neighbouring southeast Lancashire. Plotting a heat-map of meetings reported in the database shows that the industrial towns in the Leeds to Manchester corridor, to a lesser extent in the West and East Midlands, and more particularly in London, were well represented in the coverage of advertised meetings. The strength of London reporting was unexpected. The *Northern Star* coverage of other areas was much weaker, and therefore scholars should compare reportage of meetings in other newspapers to glean the wider coverage of the movement across the country. Indeed, Gwent Archives is currently conducting a crowd-sourcing project to digitize and transcribe the Chartist newspaper *Western Vindicator*, which will provide valuable comparative material to fill this gap in our knowledge about Welsh Chartist meetings.[36]

The project we have documented here involved a carefully planned workflow: acquiring, cleaning, geocoding, and presenting hundreds of meetings extracted from millions of words of mutable newspaper text. While this workflow allowed Navickas to understand Chartism better, it has the potential to help historians identify sets of relevant texts from within any wider corpora and transform them into mappable entities that can be shared as historical datasets or visualized and interpreted. This article shares that workflow with the hope that it will facilitate the development of more historical datasets and a broader sharing of methods in historical research.

36. 'Unlocking the Chartist Trials', <http://chartist.cynefin.wales/transcribe> [accessed 20 September 2016].

18

Scholars who study texts increasingly turn to computational analyses, be they based in linguistics, geography, or otherwise, and so there is a growing need to understand exactly what has been done to a set of records to produce a result. This is important not just to ensure quality and academic rigor, but also to spread these new workflows to scholars working on other time periods or places, and to stimulate responsible experimentation. By encouraging the documentation of workflows, we can put computers to work for us, so that we can pursue our real interests, which are answers to humanities questions.

Katrina Navickas

University of Hertfordshire

k.navickas@herts.ac.uk

Adam Crymble

University of Hertfordshire

a.crymble@herts.ac.uk