

Digital histories: brief lecture notes – fill in your own notes and questions too

Workshop 1: introduction to a couple of key concepts in digital history, and a couple of challenges

Part 1: digitization of primary and secondary sources, and the 'googlerisation' of data

Searching before digitisation

Process:

1. visit library, county record office, or National Archives
2. search card catalogue or index files (secondary sources, e.g. books, organised by Dewey catalogue system; primary sources, e.g. archives, newspapers, organised by the particular system of the repository). These catalogues gradually transferred to computer catalogue, but system is the same.
3. If not on open shelves, order up the book/archive. This may take several hours or even days.
4. Look at the physical copy of the book/archive/newspaper, and make notes in pencil on paper, organised in the system developed by the library/archive. Some archives are microfilmed- ruin eyesight trying to decipher the microfilm on an ancient machine reader.
5. Go home and try to decipher own notes.

Process after digitisation:

1. go to computer. No need to travel anywhere to find archives.
2. Search catalogue of database using keywords. This may be organised along the system of its predecessor before digitisation, or may use an entirely new system (or no system – e.g. flickr collection of images).

3. keywords bring up selection of documents. There may be contextual information or links to similar or connected documents – Connected Histories is a good example of this.

4. download 'relevant' documents. I may or may not transcribe them. I might annotate the files.

Review of the major digitisation initiatives:

The British Library have just digitised over a million images in their book collection. They have uploaded them to flickr:

<http://www.flickr.com/photos/britishlibrary>

Read their blog about the project here:

<http://britishlibrary.typepad.co.uk/digital-scholarship/2013/12/a-million-first-steps.html>

1. choose an image from their collection.
2. why have they digitised the images? reflecting on what you read in Cohen and Rosenzweig, discuss the advantages and disadvantages.
3. compare with other institutional digitisation projects - e.g. British Museum:

http://www.britishmuseum.org/research/collection_online/search.aspx

The problems of Optical Character Recognition (OCR) and searching.

Here is the original page image from the Trial of John Horne Tooke for treason, 1794, scanned in google books.

Christopher Metcalf. Here is a mistake in the summons for I live at Weitham, in the county of Essex.—Excused.

John Thompson, Thomas Lewes, and Nathaniel Allen, each severally declared they were not freeholders, and were excused.

Edward Hill. I am a freeholder of ten pounds a-year.

Mr. Erskine. I challenge him.

Sam Farmer. I am very much afflicted with the gout, and I have it on me now.—Excused.

James Crumpton. i labour under a disorder in my bowels, which renders it impossible for me to undergo confinement any length of time, and I am confident, I shall not be able to fulfil the duties required.—Excused.

John Lovett and John Peavey were dismissed, as not being freeholders.

Thomas Gildart. I am a freeholder.

Mr. Erskine. I challenge him.

The Clerk now informed the Court, that he had gone through the pannel, and there were but nine Jury men sworn: when the Court ordered the defaulters to be called on their juries, which was proceeded in, until they came to th: name of le y mier, who now answered to his name, but it was observed, he had been put by on account of his ill health. Whereupon,

Here is the full text deciphered by OCR. Ring the mistakes.

Thomas Druce, Tho has Maud, James Shepnell, and Benjamin Bradbury, each declared they were not freeholders, and were severally excused.

Christopher Metcalf. Here is a mistake in th* summons for I live at VVeitiam, in the county of t*flex. — excused.

John Thompson- Thomas Lewe.?, and Nathaniel Allen, each severally declared they w.:re not freeholders, and were excused.

Edward Hill. I am a freeholder of ten pound.-, a-year.

Mr. Erskine. I chatlange hi n.

Cam Farmer. I am ve^y much a/iilSled with the gaut, and I have it on me now.—Excused.

Jarhes Crumpton. i labour under a disorder in my bowels, which renders it imposible for me to undergo confine men any length of ti.ne, and I .in confident, I shall no. be able to sulfil the duties required.—Excused.

John Lovett and Johii Peavey were dismisssd, as not btng freeholders.

Thomas Gildart. I -a m a freeholder.

-Mr. Erskine. I challenge him.

The Clerk now insormed the Court, that he had gor>s through the pannel, and there were but nine Jury inn s./orn: wnen'tie Court ordered the desaulters to be called on their juries. . uh was proceeded in, until they ca.no to t'h : name of le ,y rner, who now answered to hir. name, b.it.it ..bierv !, lie had been put by ori account of his ill health. Whereupon,

OCR is even worse with tables and diagrams.

3. Practical review of the free tools and downloads listed on the first page of the reading list.

1. Twitter: <https://twitter.com/>

Set up your own twitter account, follow @TBC, and look at the twitter feeds of the #twitterstorians the module twitter account follows.

2. Google Advanced Search: http://www.google.co.uk/advanced_search

Make this your 'home page', and use it in preference to a normal Google Search.

Other Google services you will want to familiarise yourself with (all available from the header menu on the Google Advanced Search page):

- Blogger: Google's free blog site, available to use with a gmail account etc. Please sign up for a space to create a blog. Please feel free to use Wordpress instead: <http://wordpress.org/>
- Googledocs: A file sharing site that allows you to collaborate on creating word processing documents, spread sheets, etc.
- Google+: Google's answer to Facebook. I don't use it much, but we may want to explore it as a social medium for class discussions.

- Google Earth: We will be exploring the opportunities created by Google Earth in the session on mapping; and it would be worthwhile downloading it to your own machine. http://www.google.co.uk/intl/en_uk/earth/index.html

Other sites to evaluate (if time):

Zotero: <http://www.zotero.org/>

This is the best 'citation management' site on the web, and allows you to collect both citations to books and articles, but also websites, as you work online (creating bibliographies you can then export to an essay for example). It works best as a 'plug-in' in a browser such as Firefox, but can also be downloaded as a 'standalone' application for use with Internet Explorer.

Dropbox: <http://www.dropbox.com/>

A free file sharing site that allows you to exchange large documents, and back up your local hard disk.

Flickr: <http://www.flickr.com/>

An image sharing facility – useful for locating the right illustration, and sharing large image files.