What's a 'Liberal' newspaper anyway?

Historical and (meta)data-driven approaches to categorizing the Victorian press.

Categorizing information often amounts to an essential but contested step in digital humanities research. In this paper, we discuss some issues related to the political classification of British historical newspapers. In the first instance, we rely on a traditional resource often used by media historians: Victorian newspaper directories. These directories contain annual, close-to-exhaustive lists of newspapers and periodicals that circulated on the British Isles (Gliserman 1969; Brake 2015). The directories provide elaborate profiles, detailing the days of publication, the circulation and potential audiences of a newspaper title. Also, the summary contains a self-reported political leaning. Figure 1 shows an example of a record taken from Mitchell's Press Directory from 1870.
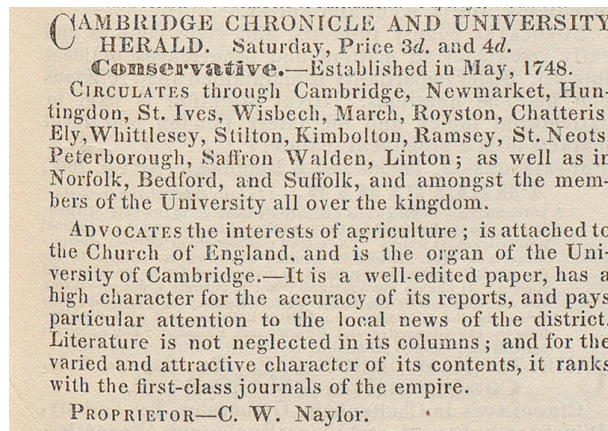

CAMBRIDGE CHRONICLE AND UNIVERSITY HERALD. Saturday, Price 3*d*. and 4*d*.
Conservative.—Established in May, 1748.
CIRCULATES through Cambridge, Newmarket, Huntingdon, St. Ives, Wisbech, March, Royston, Chatteris, Ely, Whittlesey, Stilton, Kimbolton, Ramsey, St. Neots, Peterborough, Saffron Walden, Linton; as well as in Norfolk, Bedford, and Suffolk, and amongst the members of the University all over the kingdom.
ADVOCATES the interests of agriculture; is attached to the Church of England, and is the organ of the University of Cambridge.—It is a well-edited paper, has a high character for the accuracy of its reports, and pays particular attention to the local news of the district. Literature is not neglected in its columns; and for the varied and attractive character of its contents, it ranks with the first-class journals of the empire.
PROPRIETOR—C. W. Naylor.

Figure 1: Example of the Cambridge Chronicle and University Herald *as recorded in Mitchell's* Newspaper Press Directories *of 1870*

Recently, the Living with Machines project[1] digitized and processed Mitchell's Newspaper Paper Press Directories (Gliserman 1969), and converted the scans to structured data. Figure 2 shows the output of this process. After automatic processing, the directories have been manually checked and corrected.

TITLE	POLITICS	DISTRICT	COUNTY	PRICE	DATE	DESCRIPTION	LINK_ID
ALTRINCHAM AND BOWDON GUARDIAN	neutral	Bowdon	Cheshire	2d; 3d	1862	Takes no part in politics , […].	null
BRACKLEY OBSERVER	liberal	Brackley	Buckinghamshire	1½d	1858	Taking a liberal view of politics . […]	CID_00570

Table 1: Example output of the processing pipeline for the year 1883

Historians have made primarily 'extractive' use of these directories, treating them as resources to harvest facts and information about historical newspapers. However, as (O'Maley, 2015) argues, directories should be treated and investigated as historical sources in their own right. O'Maley discusses Mitchell in detail, demonstrating how these directories—and the typology they provided—did not just reflect but also shaped the historical newspaper landscape. In a similar vein (Brake, 2015), observes: "Annual listings in the press directories are of immense value to scholars: not only do they provide a trajectory of prices hard to find elsewhere except through examinations of issues, but they also represent the industry diachronically and synchronically, offering information about changes in titles, readerships, publishers, illustrations, and geographical distribution." In this presentation, we investigate the politics of the press via the lens of these historical directories. In this sense, we leverage "vintage" classifications as an entry point to understanding what labels such as "liberal" or "conservative" actually meant when applied to the Victorian press.

Firstly we inspect the **distribution** of political labels in Mitchell. We observe that while a few categories clearly dominated—being very frequent—the political typology was very extensive and detailed. The directories contain a long tail of labels, allowing for finer nuances to shine through. In total, we recorded a total of 70 political labels (after normalization, i.e. removing all the OCR errors).
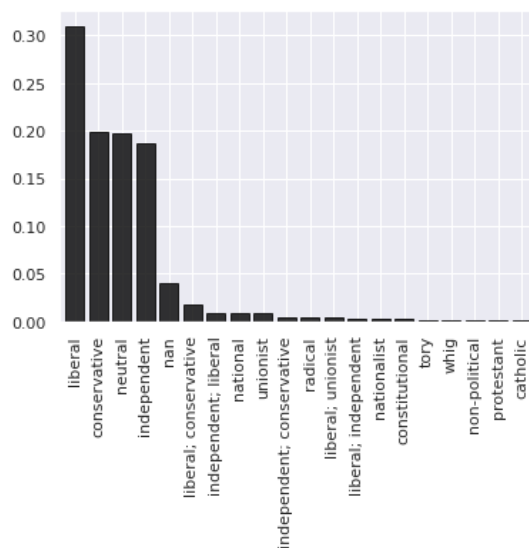

Figure 2: Distribution of political labels

Secondly, we calculate the frequency of **label transitions**. We linked the annual records of newspaper titles over time, which allows us to study their diachronic evolution, and then computed the extent to which newspapers changed their orientation. Table 2 shows the ten most frequent transitions.

source	target	count
neutral	liberal	127
neutral	independent	113
independent	liberal	108
liberal	independent	105
neutral	conservative	67
liberal & conservative	conservative	43
conservative	independent	40
conservative	liberal & conservative	29
liberal	conservative	29
independent	unionist	26

Table 2: Frequency of the ten most common shifts in political orientation

Inspecting political affiliation over time allows us to investigate the use and semantics of labels. We see that political orientation is not a permanent fixture: one-quarter of the newspapers that appear more than once changed their orientation during their run (the proportion is even higher for newspapers that re-appear more than three times, running to 35%). Table 2 lists the most frequent transitions, indicating that titles rarely 'crossed the floor'—in the sense of switching directly from liberal to conservative or vice versa (only 45 in total). The neutral and independent categories appear more fluid and these labels were most susceptible to change. Not only are they more interchangeable among themselves (we noted 133 switches in total) but also with liberal and conservative titles.

A temporal reading of Mitchell's typology suggests that the political orientation of newspapers was far from set in stone, but some borders were more porous than others. The somewhat elusive category of independent and neutral periodicals emerged as playing a critical role, not only did they absorb large swaths of the more partisan press, it also was the source from which many liberal or conservative papers later emerged. Put in numbers: we noted 229 transitions of liberal/conservative to neutral/independent, and 375 in the other direction. In other scenarios, the changes remain relatively modest, but nonetheless revealing. For example, many transitions between 'liberal & conservative' to conservative suggest these are interchangeable or at least similar. The conservative press seems more resilient compared to its liberal counterpart, with the latter appearing almost twice as often as the source of a transition.

Thirdly, we investigated the **relationship between the political labels and the discursive representations** of newspapers in the directories. We estimated the association between metadata and topic prevalence using a Structural Topic Model (STM). Figure 3 shows the topics positioned by the strength of their association with liberal or conservative.
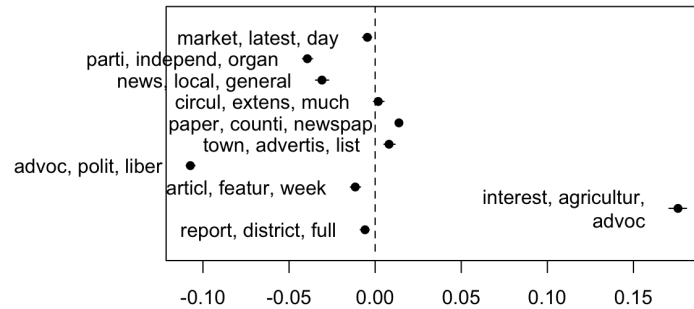

Figure 3: topics positioned by the strength of their association with liberal (left) or conservative (right).

By incorporating metadata in the model, STM enables us to infer the strength of association between topics and document-level variables. Figure 3 plots the change in topic proportions shifting from liberal (left) to conservative (right). Interestingly, strong "ideological" effects are largely absent, i.e. most of the topics in newspaper descriptions are centered around zero, meaning that the content seems largely neutral with respect to the political orientation of the newspapers. Topic 'advoc', 'polit' and 'liber' tilts to liberal, but only minimally so. The effect size is significant but small, and the same applies to most other topics. The fact that so few topics clearly aligned with party orientation, indicates that Mitchell refrained from further emphasizing ideological distinctions in his periodical descriptions. Put differently: the *Directories* applied the same topics to both liberal and conservative papers, foregrounding aspects such as the scope and circulation of a paper, the interests it promoted or other characteristics (such as connections by rail or telegraph) instead of political features. In the presentation, we will argue that the directories increasingly downplay the political aspects of the press. Even though Mitchell continues to label titles by their political orientation, from a discursive point of view, the portraits of liberal and conservative papers become indistinguishable.

Up to this point, we investigated the newspaper landscape only through the angle of vernacular metadata. In future work, we focus on the **relationship between newspaper content and metadata**. To do this, we use a novel (and hopefully soon to be open-access) corpus of "newspaper ngrams" (Michel et al, 2011; Pechenick et al, 2015). For each digital newspaper (including all newspapers from the British Newspaper Archive and other collections such as JISC and Heritage Made Digital) we collected monthly uni- and bigram counts and linked those to the historical metadata derived from the directories. In our ongoing analysis, we follow the supervised approach of (Peterson & Spirling, 2018), who use classifier accuracy as a measure of ideological disagreement. By doing this we can establish the extent to which patterns in the textual content can be used to predict the political orientation of newspapers.

References

Brake, L., 2015. Nineteenth-Century Newspaper Press Directories: The National Gallery of the British Press. Vic. Period. Rev. 48, 569–590. https://doi.org/10.1353/vpr.2015.0055

Gliserman, S., 1969. Mitchell's "Newspaper Press Directory": 1846-1907 21.

O'Malley, T., 2015. Mitchell's Newspaper Press Directory and the Late Victorian and Early Twentieth-Century Press. Vic. Period. Rev. 48, 591–606. https://doi.org/10.1353/vpr.2015.0057

Michel, J.-B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Team, G.B., Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., others, 2011. Quantitative analysis of culture using millions of digitized books. science 331, 176–182.

Pechenick, E.A., Danforth, C.M., Dodds, P.S., 2015. Characterizing the Google Books Corpus: Strong Limits to Inferences of Socio-Cultural and Linguistic Evolution. PLOS ONE 10, e0137041. https://doi.org/10.1371/journal.pone.0137041

Peterson, A., & Spirling, A. (2018). Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems. Political Analysis, 26(1), 120-128. doi:10.1017/pan.2017.39