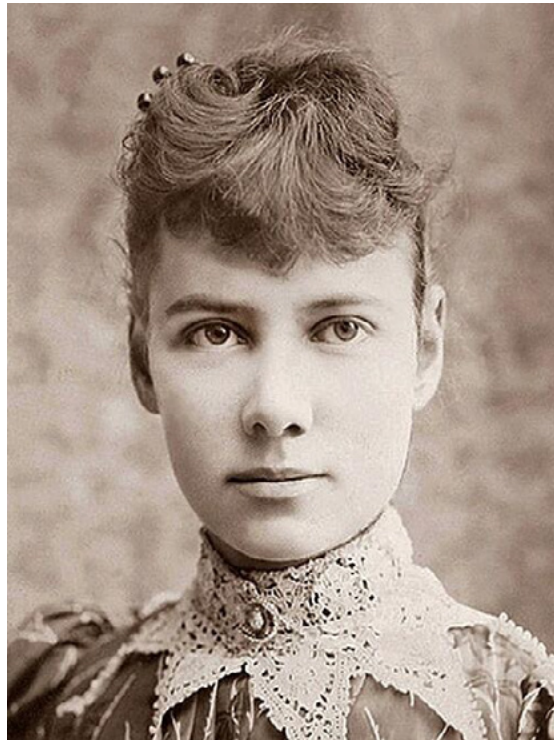




Nellie Bly, pioneering female stunt reporter.

Female Stunt Reporters

By Robert Lawson – Birmingham City University, Maryum Firdous – Coventry University, and Ashley Barlow – Coventry University

Summary

This project investigates how female stunt reporters and female journalists are represented in a cross-section of US and UK press outputs, focusing on 1880-1920. Our analysis suggests a generally negative sentiment surrounding women working in journalism, echoing broader sexist discourses at the time.

Introduction

Female stunt reporters (FSRs) were pioneering investigative journalists in the late nineteenth century who exposed social injustices. Their work led to tangible changes, such as increased funding for psychiatric care and the establishment of San Francisco's first ambulance service. We chose this topic as we are interested in seeing how gender norms were challenged and how these women tried to transform society a century ago.

Research question: How FSRs and female journalists were represented in a cross-section of US and UK press outputs.

Methodology

Data

We started with the pre-made data set available in *Gale Digital Scholar Lab* called 'Female Stunt Reporters', but this contained only 91 documents, which proved to be insufficient for detailed analysis. We then created a number of sub-corpora, as follows:

US corpus

- "Female journalist": articles 1880-1920 (56 documents)
- "Female journalist": editorials 1880 - present day (200 documents)
- "Lady journalist": articles 1880-1920 (99 documents)

UK corpus

- "Female journalist": articles 1880-1920 (39 documents)
- "Female journalist": editorials 1880 - present day (915 documents)
- "Lady journalist": articles 1880-1920 (100 documents)

Analytical tools used

We started our analysis by attempting to identify the main figures represented in our data using the 'Named Entity Recognition' tool. Although this was motivated by a desire to home in on specific outputs where FSRs were mentioned, it proved unsuccessful due to the inclusion in our data set of a fiction 'mystery detective' book which referred to multiple characters. This approach was abandoned when the highest frequency term returned in the analysis was the name 'Margaret', which wasn't especially useful. Moving forward, we used the following analytical tools:

- Ngrams for frequencies of words/phrases
- Sentiment analysis for temporal changes in attitudes towards female journalists
- Topic modelling to determine which topics were discussed in relation to female journalists

Results

1. Top 25 ngrams: Female journalist (UK batch - to 1920) = Titles (Mr, Miss, Mrs), gender referents (women, men), evaluative language (good, great)
2. Top 25 ngrams: Female journalist (US batch; editorials) = Titles (Mr, Sir – *but no Mrs or Miss*), gendered reference (man, men)
3. Top 25 ngrams: Female journalist (UK batch; editorials) = Titles (formal Lady, Mr, Miss, Mrs), lady as modifier (lady journalist), locations (London)

"A lady war correspondent! The idea was too comic. We could not believe it."

Topic modelling:

LADY n5 journalist (UK batch)



Topic Modelling Analysis

Through the topic modelling analysis, certain recurring themes are apparent. Topic 8 could be interpreted in line with societal gender specific perspectives that place 'beauty and fashion' within the sphere of female journalism. Such biases are widely observed in the documents of the data set, where the language often reduces the professionalism of female journalists to their appearance, irrespective of the assignment that they are undertaking.

Political themes that are noted perhaps reflect the shifting political situation in relation to women's suffrage in the USA, alongside the rise of female journalists operating in this sphere. Whilst nothing specifically indicating equal rights for women is visible in the topic modelling, it may suggest a highly politicized context which the documents are reporting on, and the wider nature of journalism of the period.

Certain journalists, such as Elizabeth Banks, use their agency to promote suffrage movements through their writing, either through investigative journalism or through written analogies, which in turn have influenced the outcomes of the topic modelling.

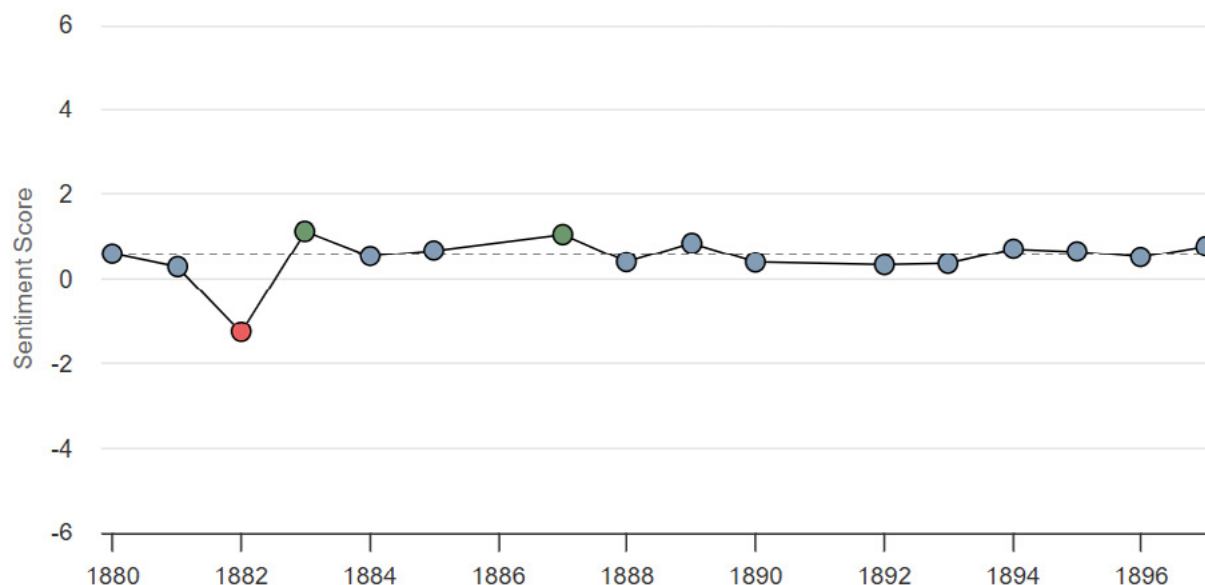
Average Word Length

The average length, in characters, of the top words. MALLET suggests that longer words often have more specific meaning and therefore topics with a higher average word length may be more specific.



Attitudes Towards Female Journalists

Through the sentiment analysis tool, it can be observed that most references from the data set were categorized as neutral, with occasional outliers that tended to reflect positive attitudes. On closer inspection, these outliers reflected articles about the journalists themselves, with particular attention paid to female American journalists operating in England, liaising with successful members of American society. Reporters singled out include Elizabeth Banks, Lydia Pinkham, and Betsy O'Dowd for their passion and innovative approaches to journalism. The negative outlier reflects challenges related to the contextual associations with the software, and the reference to Mount Vesuvius has inadvertently been included in the analysis.

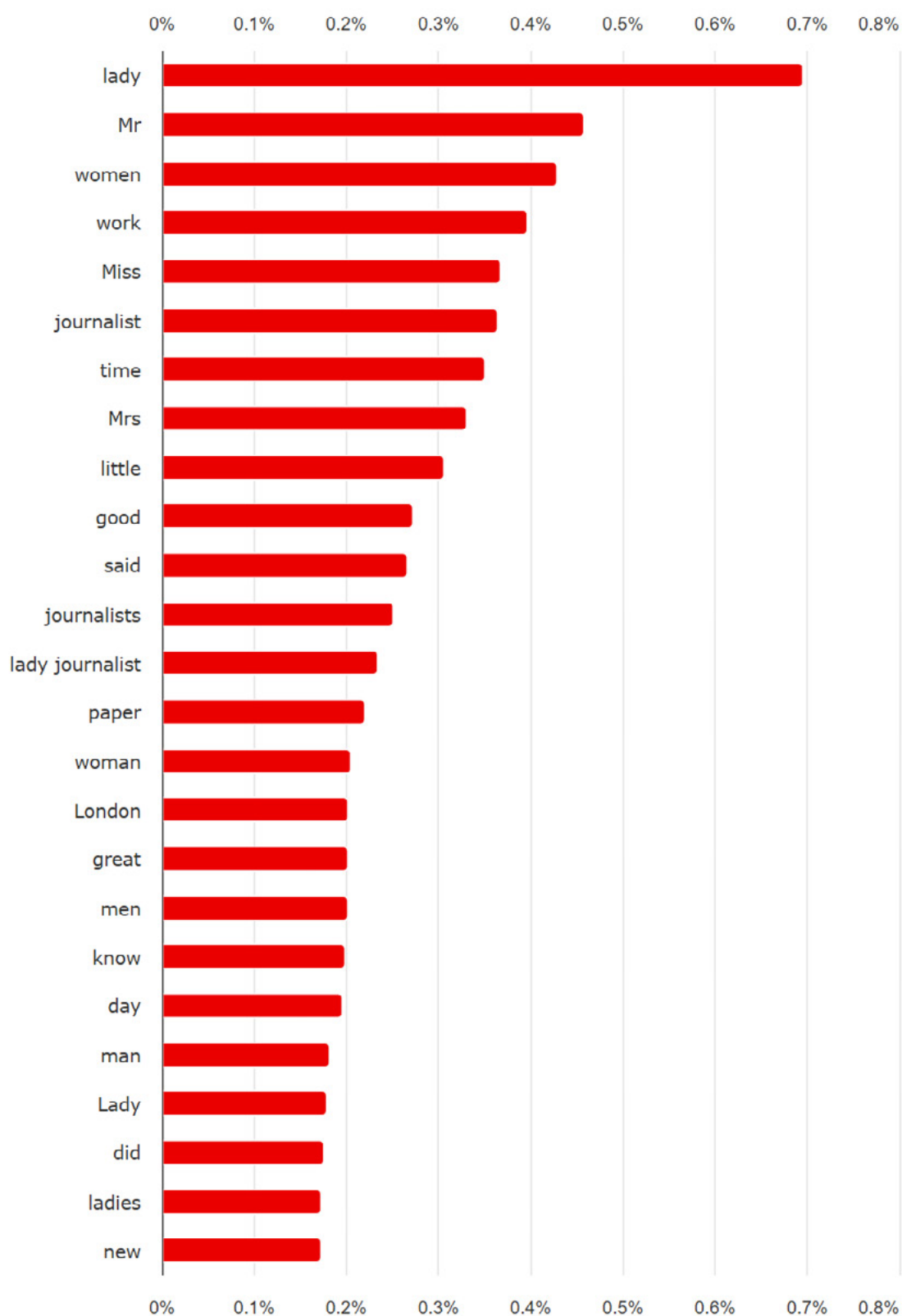


Female Journalists vs. Lady Journalists

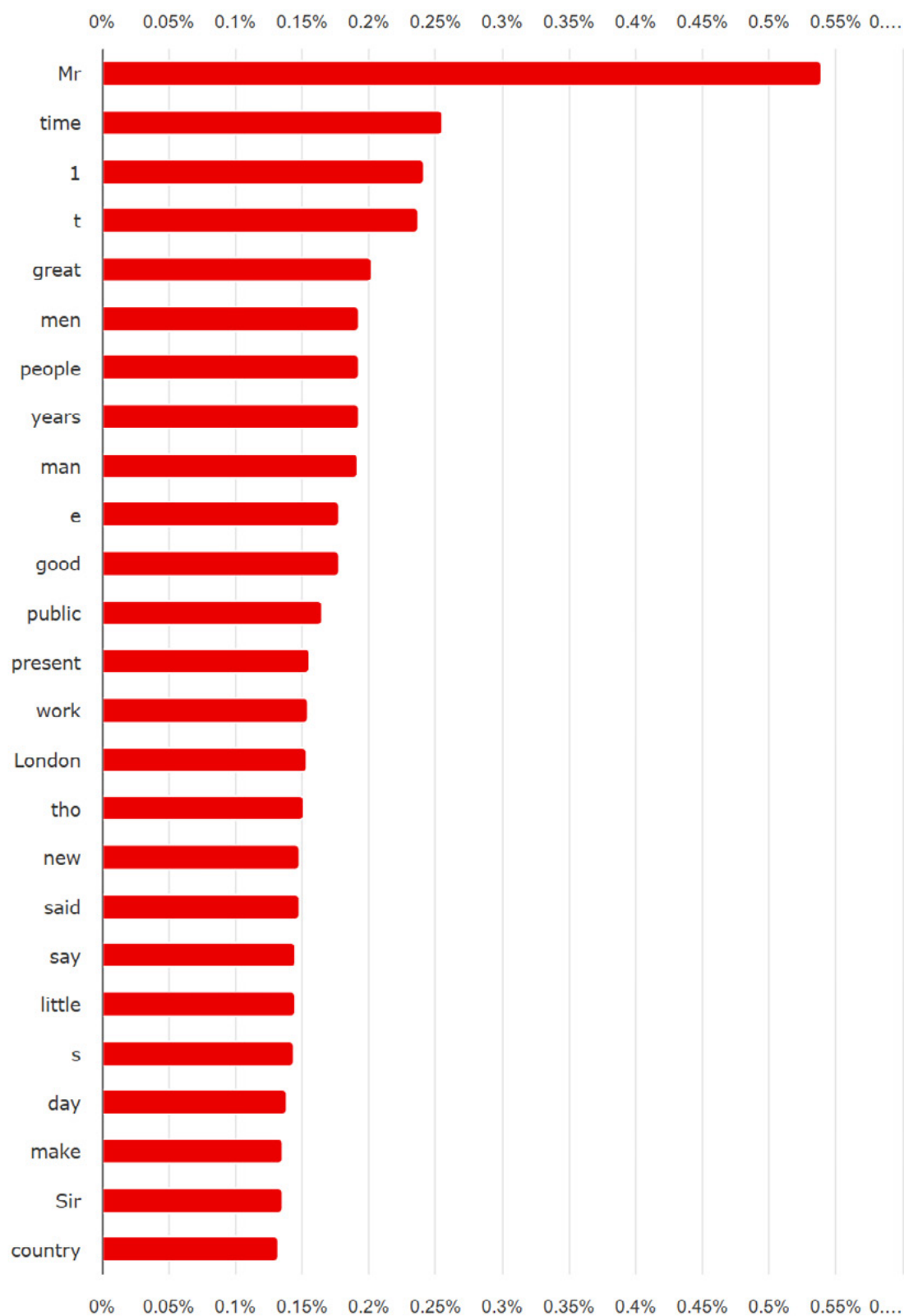
The term 'female journalist' appears 16 times across the data set. Although this is not a hugely significant representation, it is more representative than the 0 cases of 'lady journalist.' The presumption that such a job title would be well represented in the data set was largely based on social conventions of occupational titles from the late nineteenth and early twentieth centuries. The preferred use of the term 'female journalist' may reflect a recent increase in the number of women working in the journalism field, and the rising agency of women's suffrage movements in the latter decades of the 1800s.



LADY n5 journalist (UK batch)



Female journalist (UK batch; editorials)



Reflections: Trials, tribulations, and turmoil

Reflections on data...

Our initial data set was too small (<100 documents) for close historical or linguistic analysis, highlighting the importance of considering scope in the early planning stages. We then reviewed potential avenues of enquiry, building on our interest in FSRs and expanding our focus to a broader concern of women in journalism. With this in mind, we developed a number of sub-corpora using a range of search terms, including "LADY n5 journalist" and "FEMALE n5 journalist". These searches were restricted to UK and US periodicals and editorials, but OCR reliability was not always high enough to ensure good data retrieval. On review, making clearer decisions about OCR score, date range, and output type would have helped with the early stages of data construction.

Reflections on methods...

It took time to learn how to operate the various analysis functions offered on the platform and how to move between the underlying data sets. Moving forward, clearer marking and annotation of the 'runs' conducted would help with remembering what set-ups were used, what cleaning protocols were used, and what the output was. The pairwise comparisons that were applied were also less motivated than they could have been, so having a clearer idea earlier in the project of what to compare would have been helpful.

Reflections on analysis...

To facilitate content sharing, the group used an open Google Docs form. While this was useful to brainstorm research questions and shape the overall trajectory of the project, it was less useful in sharing visualizations and graphs of the analysis. Using the *Gale Digital Scholar Lab*'s internal 'note taking' function resolved many of these issues, so spending time learning this functionality would have been a useful step.

In terms of our analysis, this tended towards the more speculative side, partly because of our unfamiliarity with the underlying data and difficulty in reading through all the outputs in the time available. This could have been addressed by focusing our analysis on higher level quantitative features, rather than attempting fine-grained qualitative interpretation.

Please note: Hacking History Event Projects are produced in a single day, during the hackathon, and may not provide fully formed or finalized research. They are published on the Gale Research Showcase to show examples of what can be done in a single day within the digital humanities, using Gale Digital Scholar Lab and any other DH tools the participants choose to use. They are published to provide inspiration to others in terms of the wide range of approaches, research questions and types of analysis possible within the digital humanities.

Gale Research Showcase is a free, open repository of student-authored digital scholarship. Designed to advance early-career research, it showcases high-quality, peer-reviewed essays that demonstrate best practices in digital scholarship.

You can use Gale Research Showcase to get inspiration and guidance on using Digital Humanities techniques in your own project – and to get published! To learn more about how to get published in Gale Research Showcase, visit: gale.com/publication-opportunity

Copyright information: Projects are published in Gale Research Showcase under a Creative Commons license, CC BY-NC-SA 4.0 DEED (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)