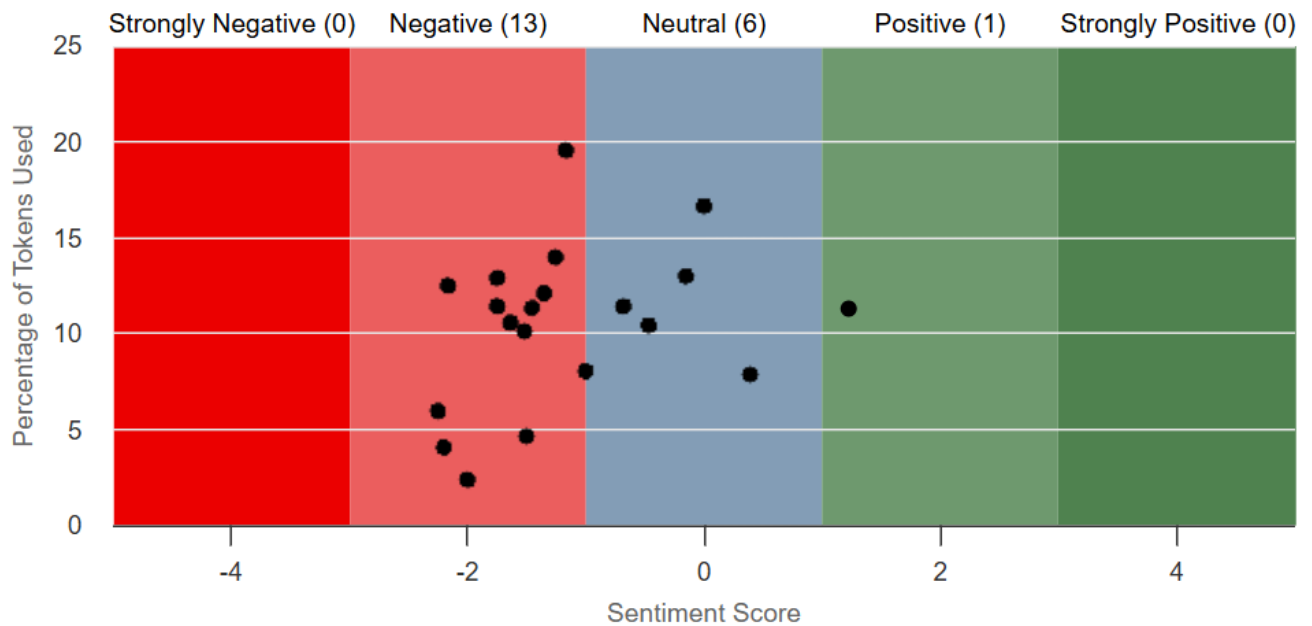




Sensation! How Has Reporting on High-Profile Murder Cases Changed Across Time?

By Simon Satchwell Giles – Birmingham Newman University, Eem Munawaro – Coventry University, Puja Lama – Coventry University, Glory Ojukwu – Coventry University

Introduction

Data and Methods

We looked at reporting on the deaths of Robert Calvi, a banker linked with the Italian Mafia and political intrigue within the Vatican City in 1982, and compared this to reporting around the death of Gareth Williams, the MI6 spy found dead locked in a holdall in London in 2010.

We looked at newspaper and other media articles gathered from *Gale Primary Sources*, focusing on reports from 15 years after the incidents we were investigating.

After experimenting with the workflow, we decided to focus on Sentiment Analysis and Topic Modelling.

A larger project may use these two stories as a starting point and bring in additional cases and data to compare.

Topic Modelling

Topic Modelling: Calvi 1

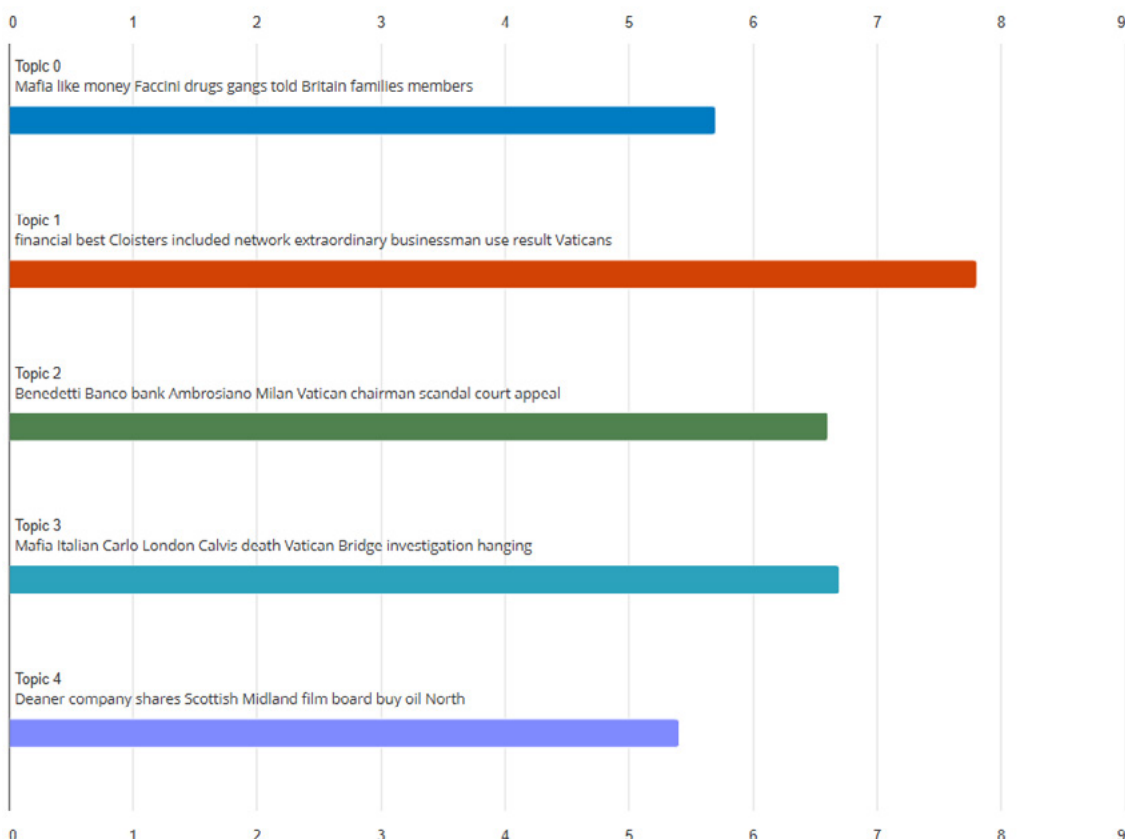
We started by analyzing 5 topic sets, to model the topics found within the Calvi data.

There were some intriguing collections of words in these topic sets, suggesting particular concerns: for example, the inclusion of 'cloisters' alongside 'network', 'financial' and 'businessman' points to a sense of secretiveness that also appears within the sentiment analysis.

However, overall, this analysis produces a confusing picture, with many words included in seemingly overlapping topics.

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.



Topic Modelling: Calvi 2

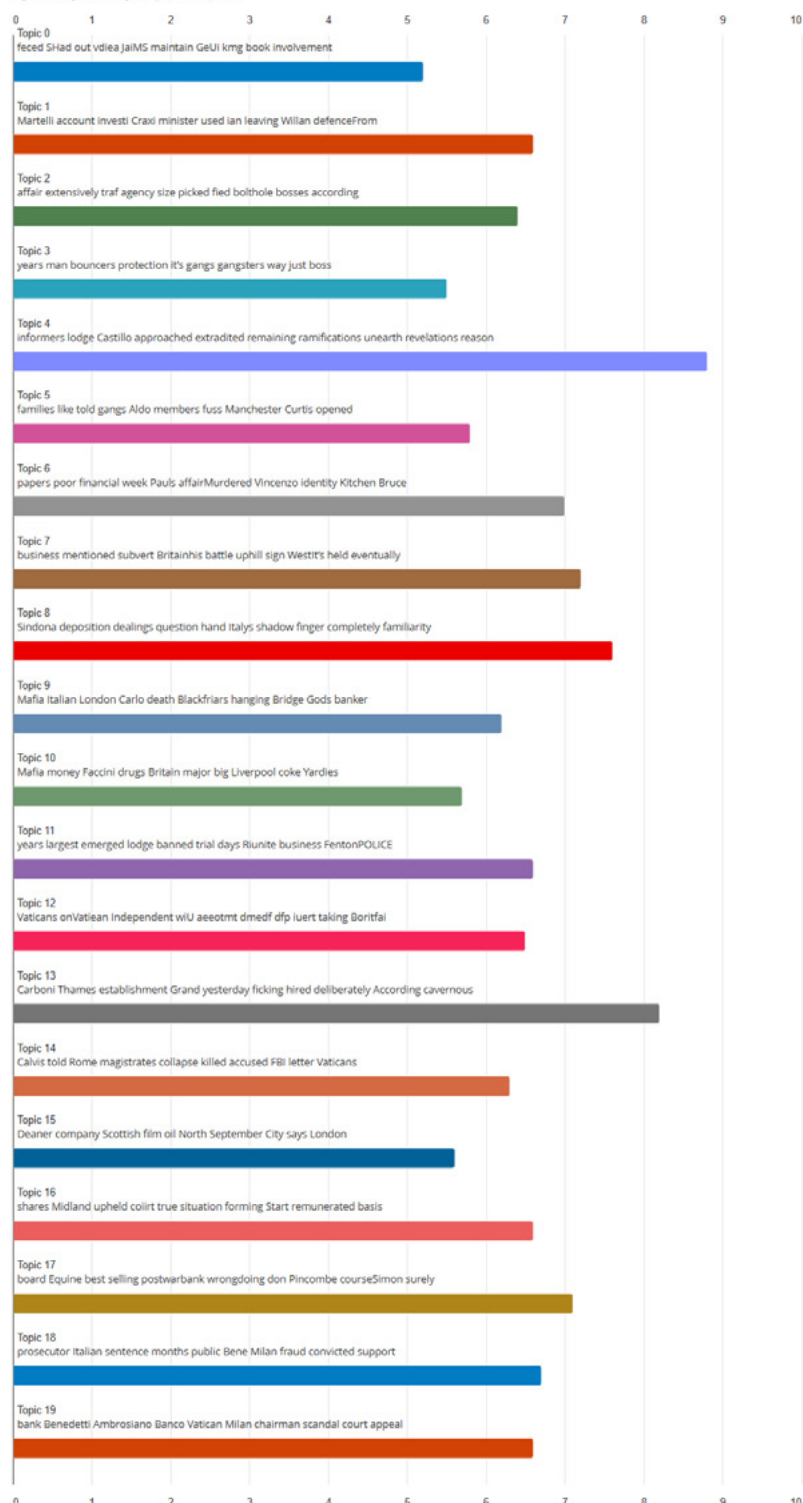
On the advice of the facilitator of the hackathon event, we ran another analysis, this time increasing the number of topics to 20, to try and draw out some more distinct categories.

This, sadly, gave us little more to go on in terms of the coherence of the topics. A potential issue here is the dataset itself; if most of the data items are short reports and updates, spread through several issues of the newspapers, it may be more difficult for these tools to effectively code the data.

When working with these datasets in the future, it may be more fruitful to use these tools as a starting point, but run them alongside traditional qualitative coding methods.

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.



Sentiment Analysis

Sentiment Analysis – Introduction

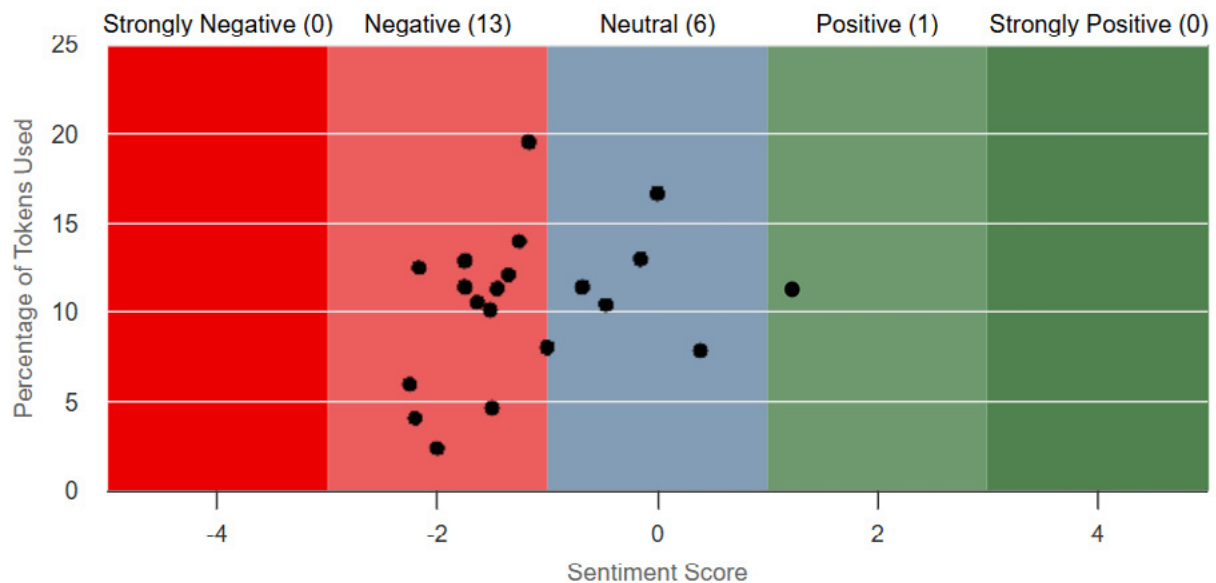
When we compared a duration of 15 years in both, in the Calvi case there were a total of only 20 documents discovered, whereas in the Williams case there were 217.

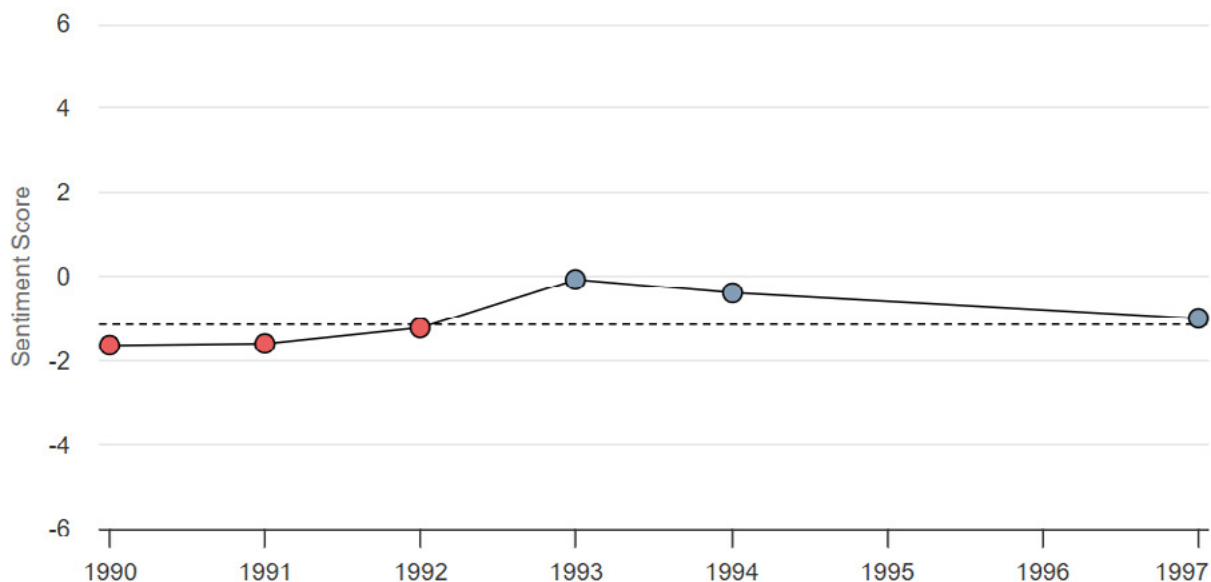
This fairly obviously represents a large difference in media culture between the two cases. In the 1980s, print newspapers were the dominant force in the media market, but there were a limited number of publications which were publishing within the limits of daily or weekly issues. By the 2010s, internet media and social media culture are in ascendance, and the number of publications as well as publication methods and constraints have been transformed.

	Calvi (1982)	Gareth (2010)
Total documents	20 (Traditional)	217 (Digital, Social Media)
Sentiment Score	-1.009	-0.819
Terms used	verdict, smuggling, foul	furious, dismissed

Sentiment Analysis - Calvi Case

A murder case, as expected, showed a generally negatively skewed sentiment analysis in the reporting.

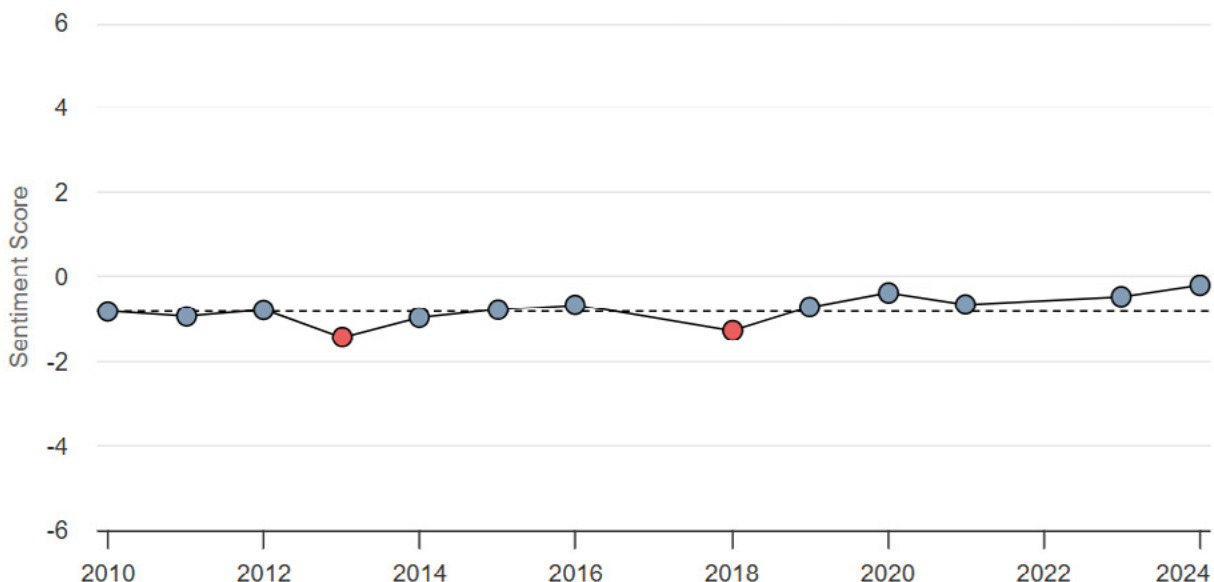




Sentiment Analysis - Williams Case

In analyzing the data of the Williams case, as expected with a tragic event, the sentiment score was skewed towards the negative, with an average of just under -0.81. However, analysis across time uncovered specific and significant dips in the reporting, first in 2013 (3 years after Williams' death) and then again in 2018.

Close analysis of the documents associated with these dips uncovered that these coincided with reporting of the conclusions of the police investigations of the inquest. A markup analysis suggested, more specifically, negative language connected with the institutions involved, including the police and criminal justice system and the security services.



Topic Modelling: Williams case

Topic Modelling in the Williams case was more successful, although the data was still noisy.

In the analysis below, several fairly coherent topics can be identified, including

- Espionage and the British intelligence services
- British establishment figures
- Gender and sexuality
- Other mysterious deaths linked to espionage and 'bungled' investigations

The Topic Modelling analysis shows promise for analysis in this project, and has also highlighted other investigations and discrete datasets that could be used in a future project.

Outcomes and Reflections

Research Outcomes

Due to time limitations, we were not able to clean the dataset as much as we wanted to. Furthermore, our research topic evolved during the time of the hackathon, and despite our wishes, we were not able to address some of the research questions that we came up with.

However, we have learnt through the process how a project might be built using these datasets that could add real insight to our understanding of changing media and shifts in public perceptions and expectations of different private, civil and cultural institutions.

Reflection

Working with Gale as our primary research tool facilitated and shaped both the process and the outcome of our project. We initially had little knowledge of how to navigate the platform, and it felt like learning on the job, however, we were able to learn by doing it consistently until we were familiar with the tools.

One of the challenges we faced was directly comparing both sources at the same time. Although Gale has strong materials on each of the sources individually, we took a longer time trying to analyze the documents individually than we would have done if we were able to compare them at once.

In the process of exploring these topics, we discovered that some articles leaned heavily on the negative side, especially in the Calvi case where conspiracy theories were widespread, however, some sources were positive and others neutral.

Overall, we had a great time exploring the tools, and discovering what we could and could not do, and it was an amazing experience we created together.

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/>)