Navigating Public Sentiment in the Circular Economy through Topic Modelling

Yingfang Yuan[1], Junhao Song[1], Zhewen Wei[1], Wei Pang[1], Jin Xuan[2], Bing Xu[3]

[1]School of Mathematical & Computer Sciences, Heriot-Watt University, [2]School of Chemistry and Chemical Engineering, University of Surrey, [3]Edinburgh Business School, Heriot-Watt University

RESEARCH BACKGROUND

To advance the circular economy (CE), it is crucial to gain insights into the evolution of public sentiments and discerning the primary concerns and cognitive pathways of the masses concerning circular chemical products, and digital technology. To achieve this, we collected data related to the circular economy from diverse platforms including **Twitter**, **Reddit**, and The **Guardian.** This comprehensive data collection spanned across three distinct strata of the public: the general public, professionals, and official sources. Next, we conducted a series of systematic experiments to gain insights into the dynamic relationship between the circular economy and public opinion. We believe our work is pioneering in investigating various levels of public opinion concerning CE through dynamic topic modelling. Meanwhile, we delve deeply into the dynamic **BERTopic** in terms of hyperparameter selection to ensure the expected performance.

RESEARCH FRAMEWORK

General Public | Professionals | Official Sources

1. Data Cleaning: Text extraction, removal of special symbols, elimination of ads, duplicates removal, language detection, etc.

2. Exploratory Data Analysis: Size, text length, word frequency analysis, etc.

3. Topic Modelling Approach Selection: Dynamic BERTopic[1] aims to analyse the evolution of topics over time. It enables you to comprehend how a subject is depicted across different temporal intervals. Meanwhile, a key advantage of BERTopic lies in the extraction of embeddings from sentence transformers. These embeddings encapsulate semantic meaning, leading to document clustering that goes beyond lexical considerations.

4. Results Analysis

5. Visualisation

6. Multidisciplinary Discussion

EXPERIMENT RESULTS – TWITTER

- We collected **Twitter** data using the **#CircularEconomy** tag, totaling 4,560 tweets between February 7th, 2023, and April 20th, 2023.
- Our analysis highlighted intriguing topics 0, 1, and 2 with the words: (waste; recycling; waste management), (packaging; recyclability), and (plastics; plastic pollution). These topics exhibited fluctuations from February to April.
- The right-side word cloud visually captures **public concerns.**

EXPERIMENT RESULTS – REDDITS

- We utilised the **subreddit** located at https://www.reddit.com/r/circular_economy/.
- This subreddit, established on July 19th, 2014, boasts over 3,000 members and ranks within the **top 10%** in terms of size.
- The dataset encompasses 708 posts.
- Our analysis revealed that topics 0, 2, and 4 are particularly intriguing for exploration, characterized by the following topic words: (waste; recycling; plastics), (circularity; systems), and (business; model). As shown in the figure above, topic 0 exhibited a consistent increase from 2017 to 2022.

EXPERIMENT RESULTS – GUARDIAN

- The **Guardian** data's experimental results are displayed in the above three figures, showcasing six identified topics, each linked to a facet of the circular economy. For instance, topic 0 highlights the rising attention to climate and emissions in Guardian content since around 2014.
- The heatmap displays correlations between topic pairs, e.g., topic 1 shows a stronger link to topic 3. The word score figure indicates word significance within each topic.
- Overall, the CE has gained more attention, especially after 2020, focusing on climate, plastics, and clothing.

SUMMARY

- Topic modelling effectively mines insights from textual data.
- Unveiling meaningful trends across diverse databases, each capturing a specific public segment's viewpoint.
- By categorising data into distinct groups related to CE, topic modelling facilitates informed investigations.
- In conclusion, we noted increased public interest in the CE. Among all the topics, **waste** and **plastics** standing out as prominent topics that attract the most public attention.

REFERENCE

[1] Grootendorst, Maarten. "BERTopic: Neural topic modeling with a class-based TF-IDF procedure." *arXiv preprint arXiv:2203.05794* (2022).