

InsightLink

● AN LLM-POWERED CONTENT ANALYSIS ASSISTANT

Members Bing-Chen Chiu, Peng-Jen Chen, Yu-Ting Chen, Ding-Hong Chen, Jeng-Yue Liu, Cheng-Shuan Lee
Advisors Dr. Chia-Yen Lee, Dr. Tzu-Fen Chang

December 2024



BACKGROUND

Recent high-profile incidents of hate crimes in the U.S., particularly those targeting Asians, have heightened concerns about their vulnerability to such violence. The onset of the **COVID-19 pandemic** has exacerbated this issue, with accusations and rhetoric labeling the novel coronavirus as an “Asian” or “Chinese” virus. This narrative has contributed to a surge in **anti-Asian hate** crimes, and even escalated to tragic events, such as the **Atlanta spa shooting** on March 16, 2021.

CASE STUDY

The Atlanta spa shootings occurred on March 16, 2021 shocked the nation, drawing attention to **the rise of anti-Asian hate crimes** during the COVID-19 pandemic. We want to explore the differences of news articles in mainstream and non-mainstream media before and after it occurred.

Comparisons Across Time

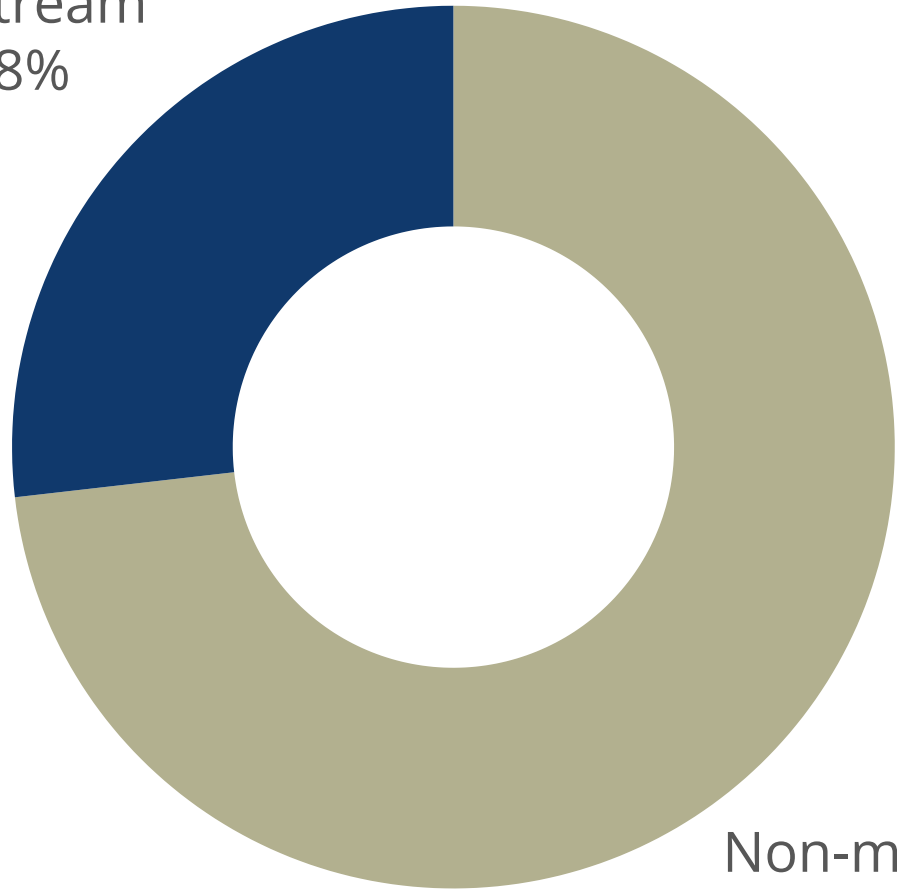
Examine whether the articles published before and after March 16, 2021 (the occurrence of Atlanta Spa Shooting) are different or similar in terms of types of racism as well as people's and organizations' reactions.

Comparisons Across Media Types

Compare the number of sources used by mainstream and non-mainstream media, as well as the types of sources they use, to determine if there are any differences or similarities.

DATASET

Mainstream
26.8%



Non-mainstream
73.2%

Number of Articles

634 (Mainstream: 170, Non-mainstream: 464)

Time Range

2020/01 - 2021/12

Publishers

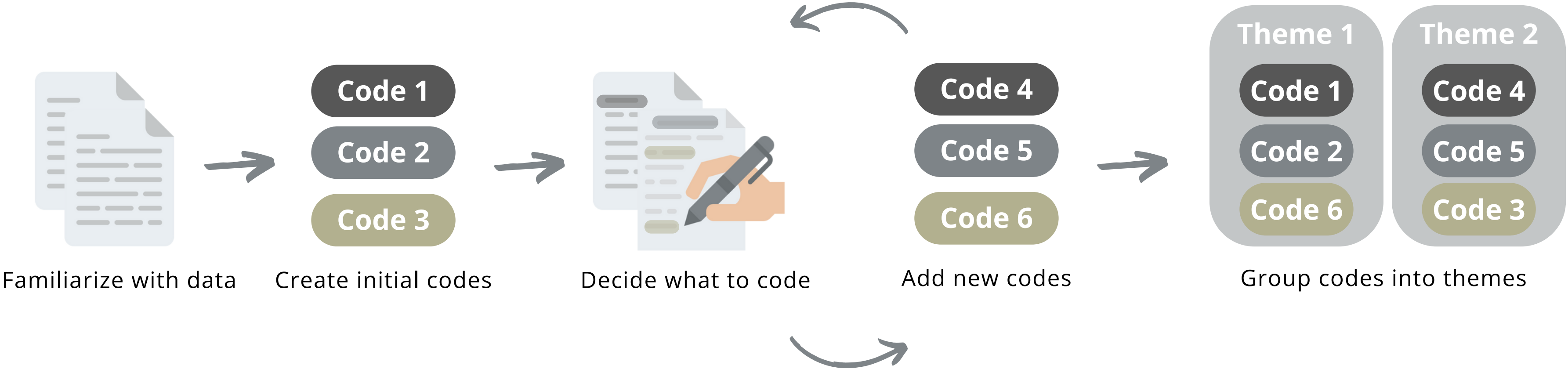
Mainstream* and Non-mainstream Media

Selected Mainstream Media

Boston Globe, LA Times, New York Times, USA Today, Washington Post, Star Tribune, Chicago Tribune

MANUAL CODING

THE TRADITIONAL APPROACH TO QUALITATIVE ANALYSIS



CHALLENGES



01 Lack of Domain Experts

Domain experts are essential for content analysis, but their specialized knowledge and demanding roles **limit availability**, and their involvement in large-scale projects adds **significant financial costs**.

02 Highly Time-consuming Process

Content analysis is a **labor-intensive** process involving the manual review and the preprocessing of unstructured data, resulting in the **difficulty in identifying key insights** effectively and efficiently.

03 Subjectivity and Bias

Human interpretations of qualitative data are inherently subjective, leading to potential bias or varying conclusions among analysts.

SOLUTIONS



- LLM-powered **content analysis assistant** with several features
- Helping users **save tons of time** of reading articles and doing content analysis manually

01

Summary of Articles

- Once an article is uploaded, **InsightLink** will analyze it.
- Users can review the summary to **understand the content more efficiently**.

02

Coding Themes & Sources with AI

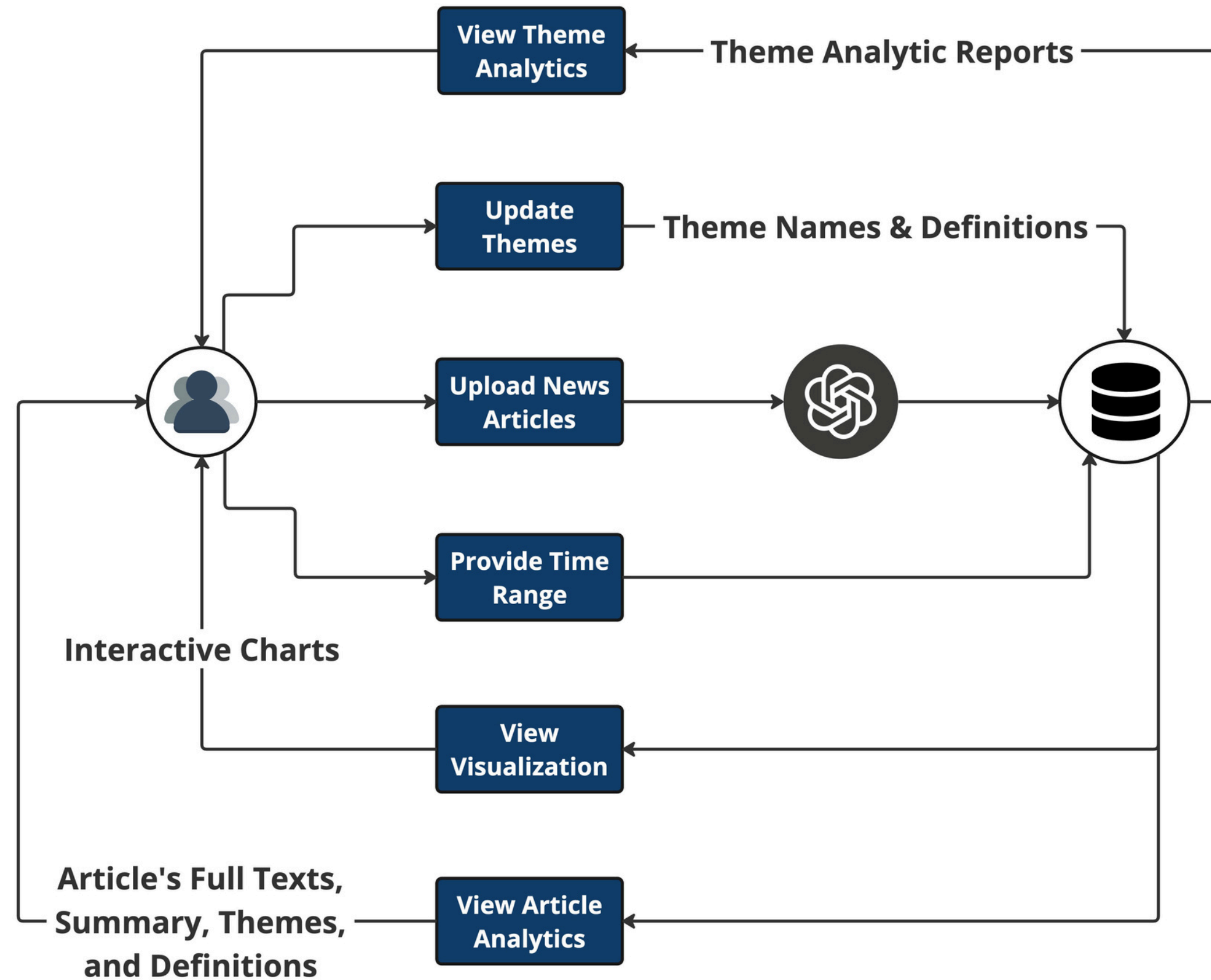
- **InsightLink** will provide **types of theme related to the main topic** (e.g., anti-Asian hate) and highlight the corresponding sentences.
- **InsightLink** will **presents the perspectives** from different people and organizations (e.g., victims, suspects and criminals, NGO).

03

Dashboards

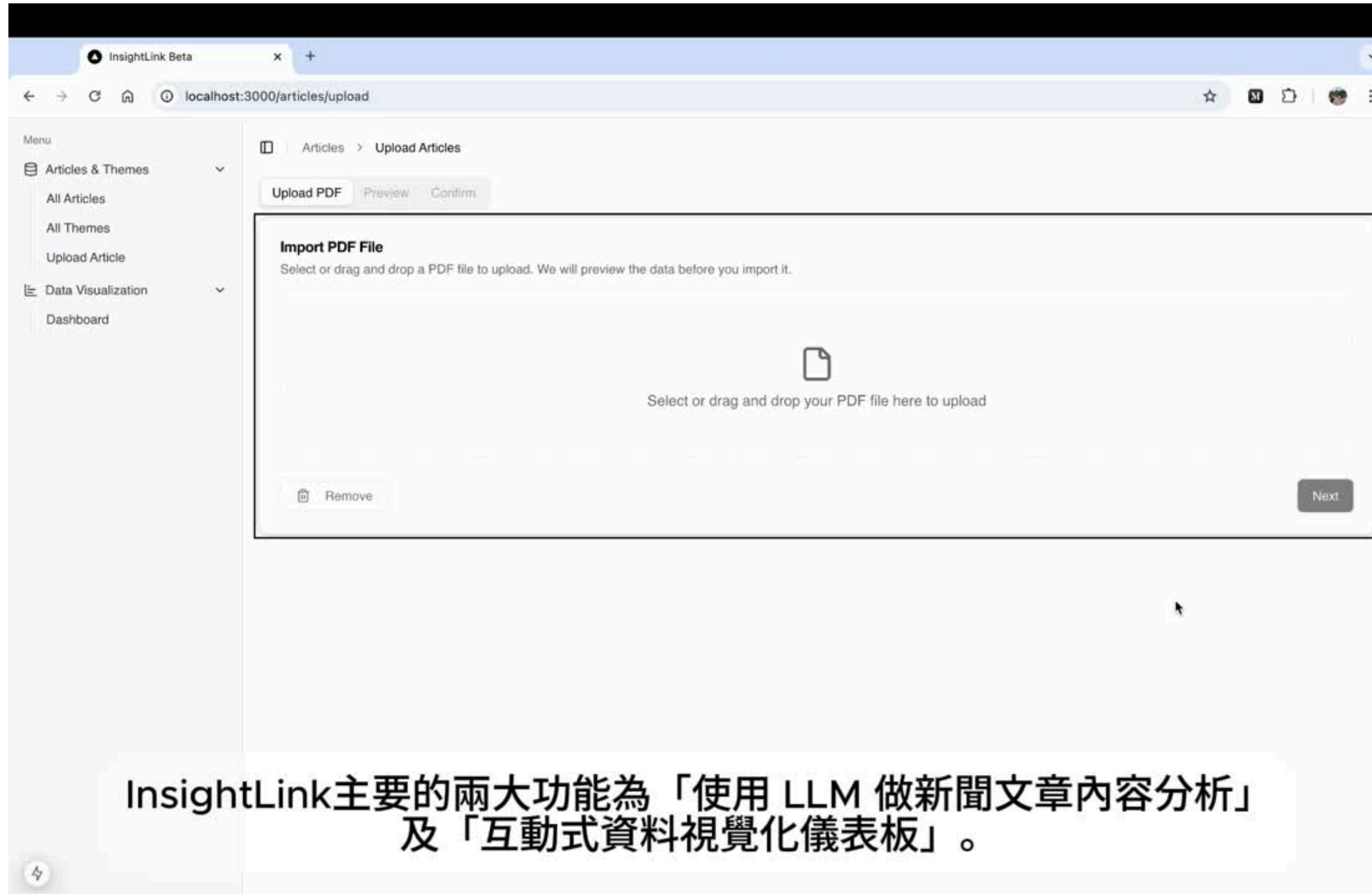
- **InsightLink** provides several **visualization of figures and tables** with statistical data.
- Users can set a specific **time interval** to understand the trend well.

DATA FLOW



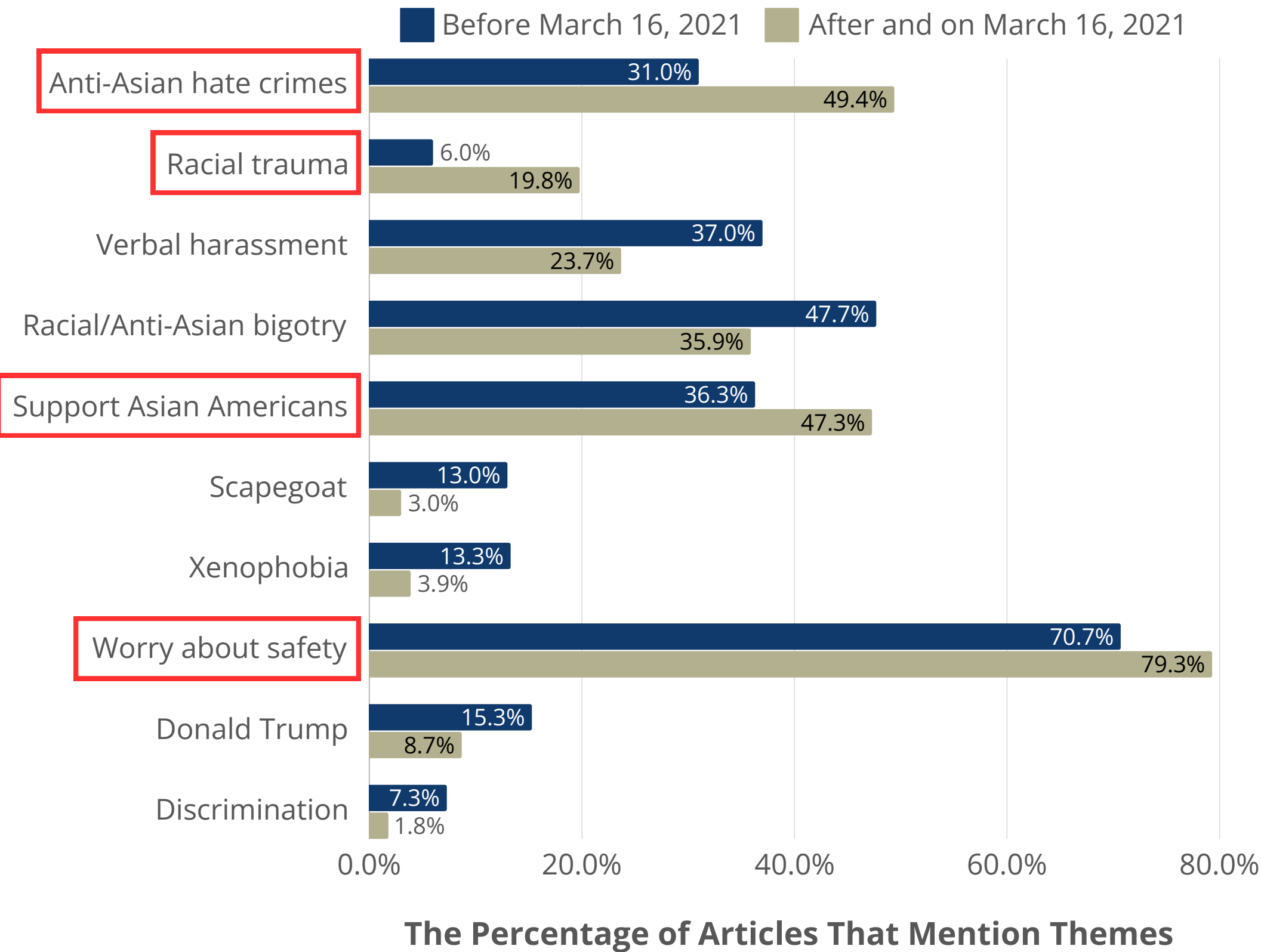
DEMO

DEMONSTRATION OF INSIGHTLINK



InsightLink主要的兩大功能為「使用 LLM 做新聞文章內容分析」
及「互動式資料視覺化儀表板」。

DATA ANALYSIS

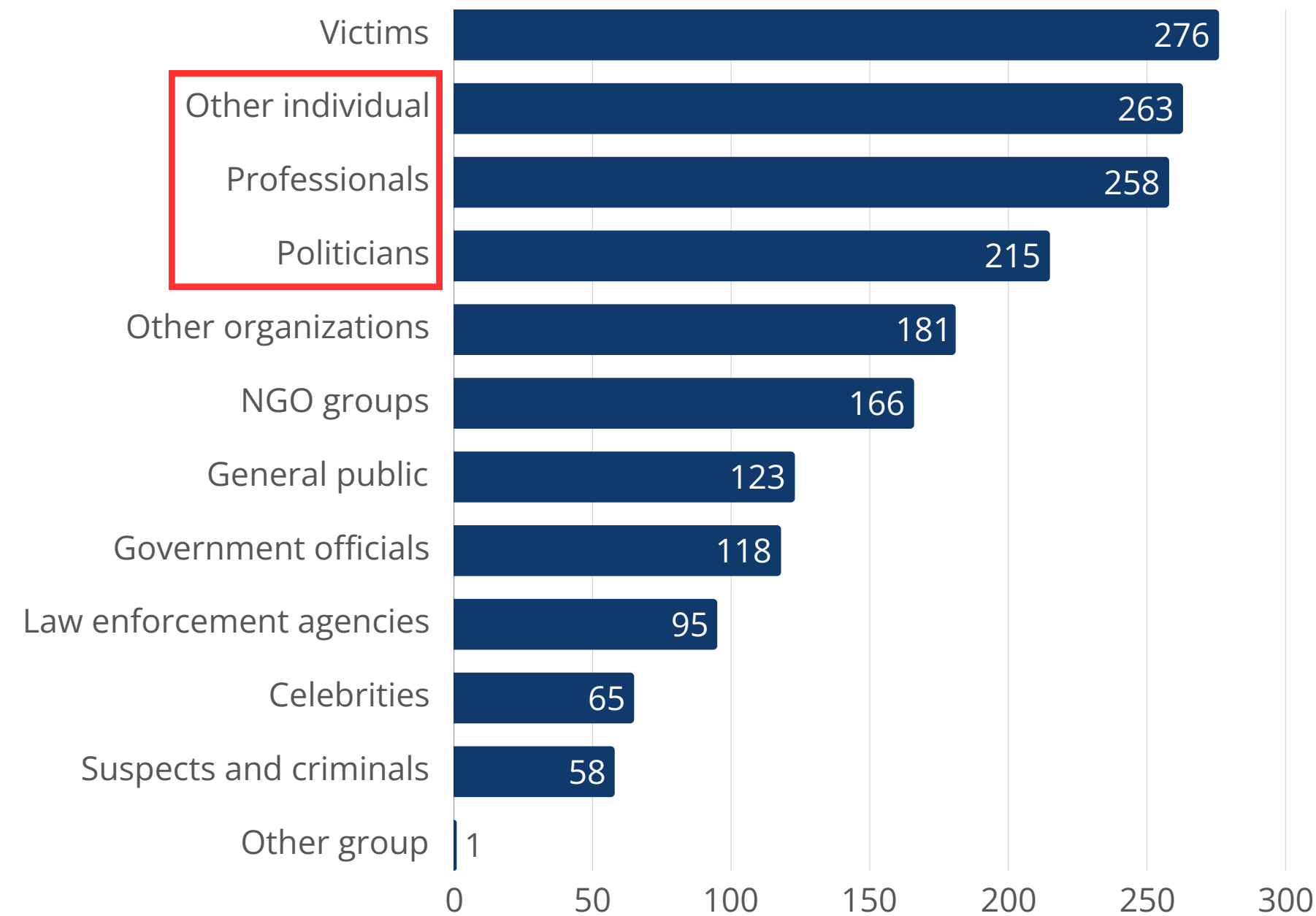


The Types of Themes Being Reported Before and After 2021/03/16

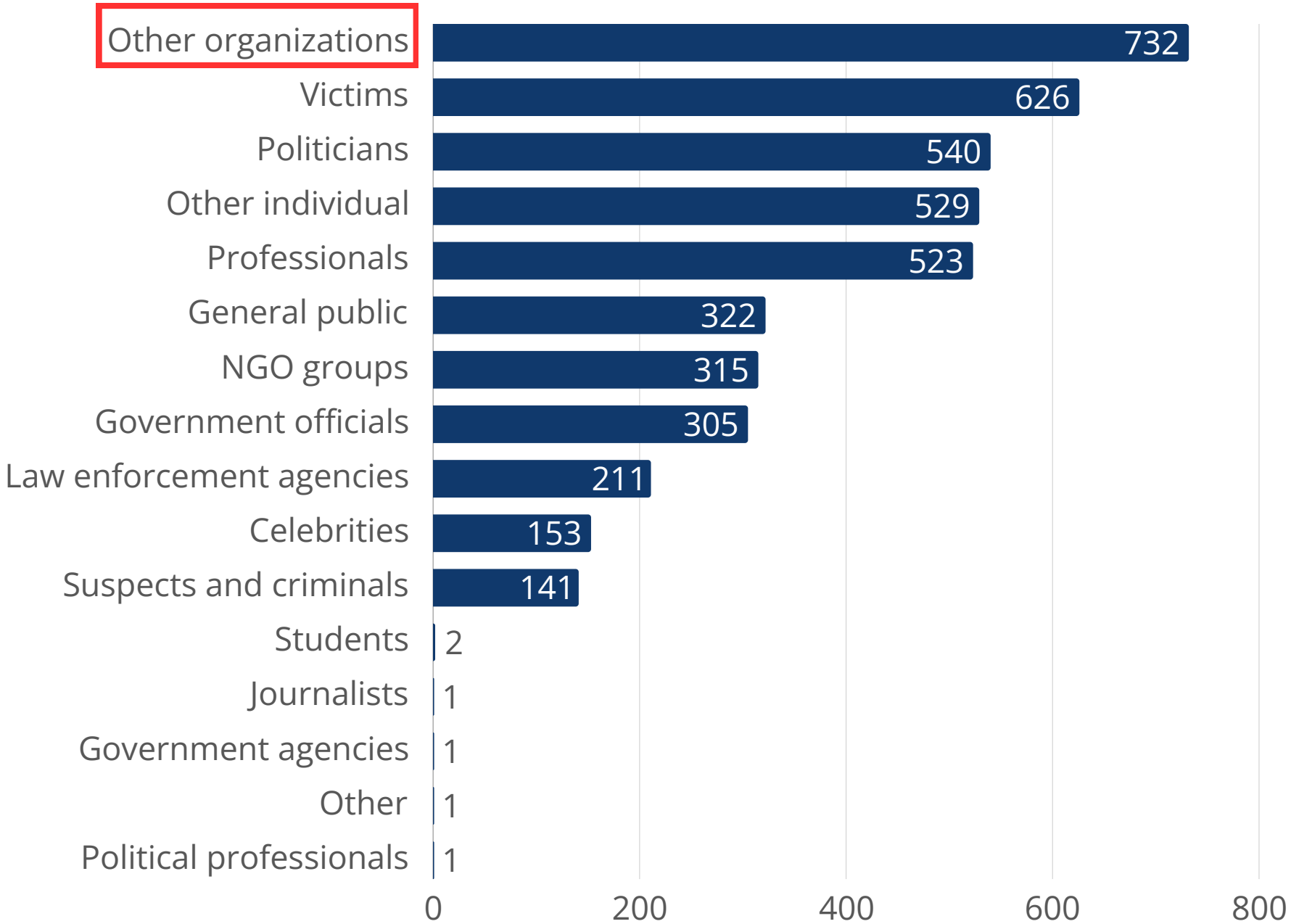
After Atlanta spa shootings, more articles mentioned **Anti-Asian hate crimes**, **racial trauma**, and **worry about safety**, which highlighted the societal reckoning and awareness sparked by significant tragic incidents.

Types of sources mostly used by the Mainstream media vs. Non-mainstream media

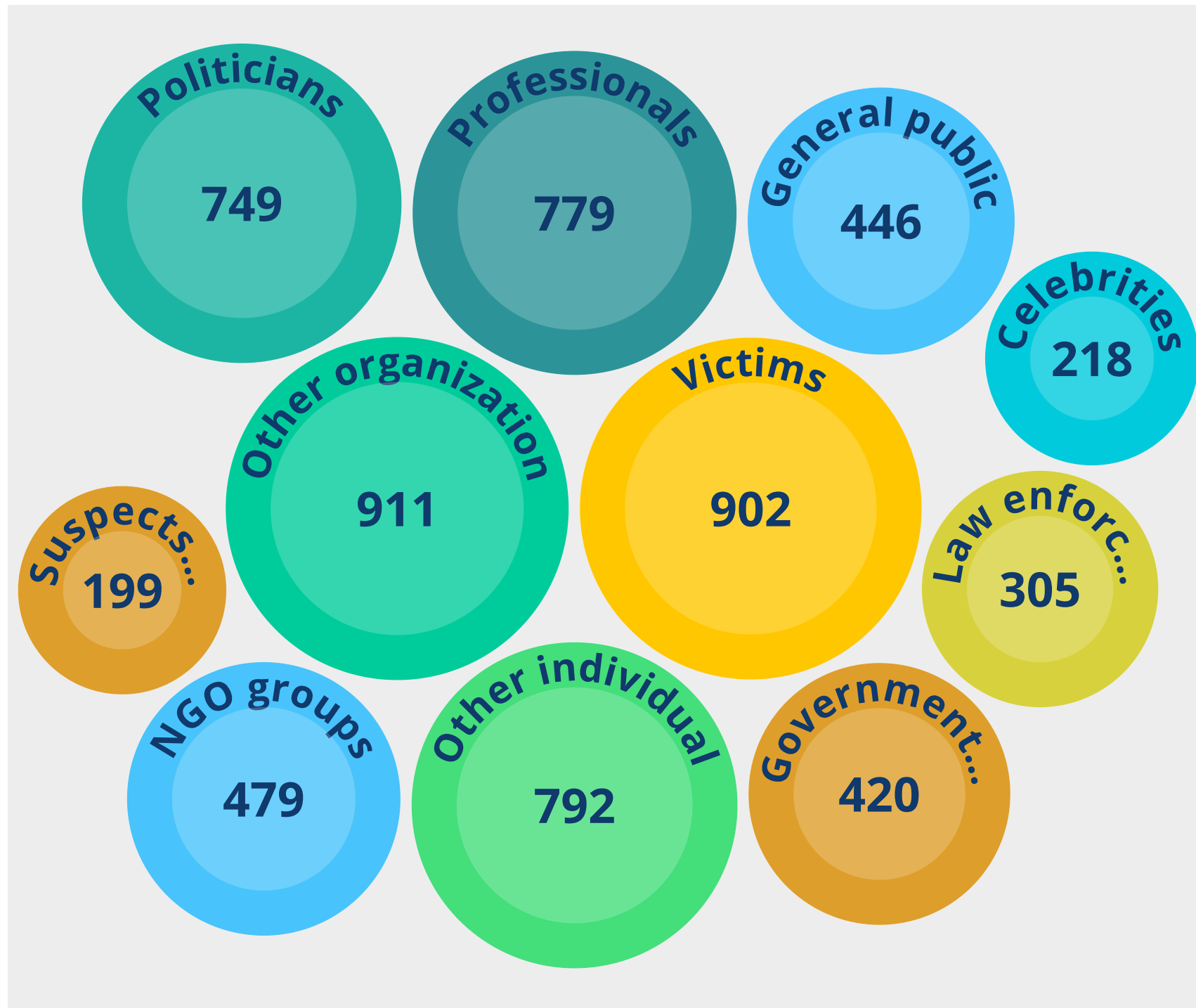
Mainstream



Non-mainstream



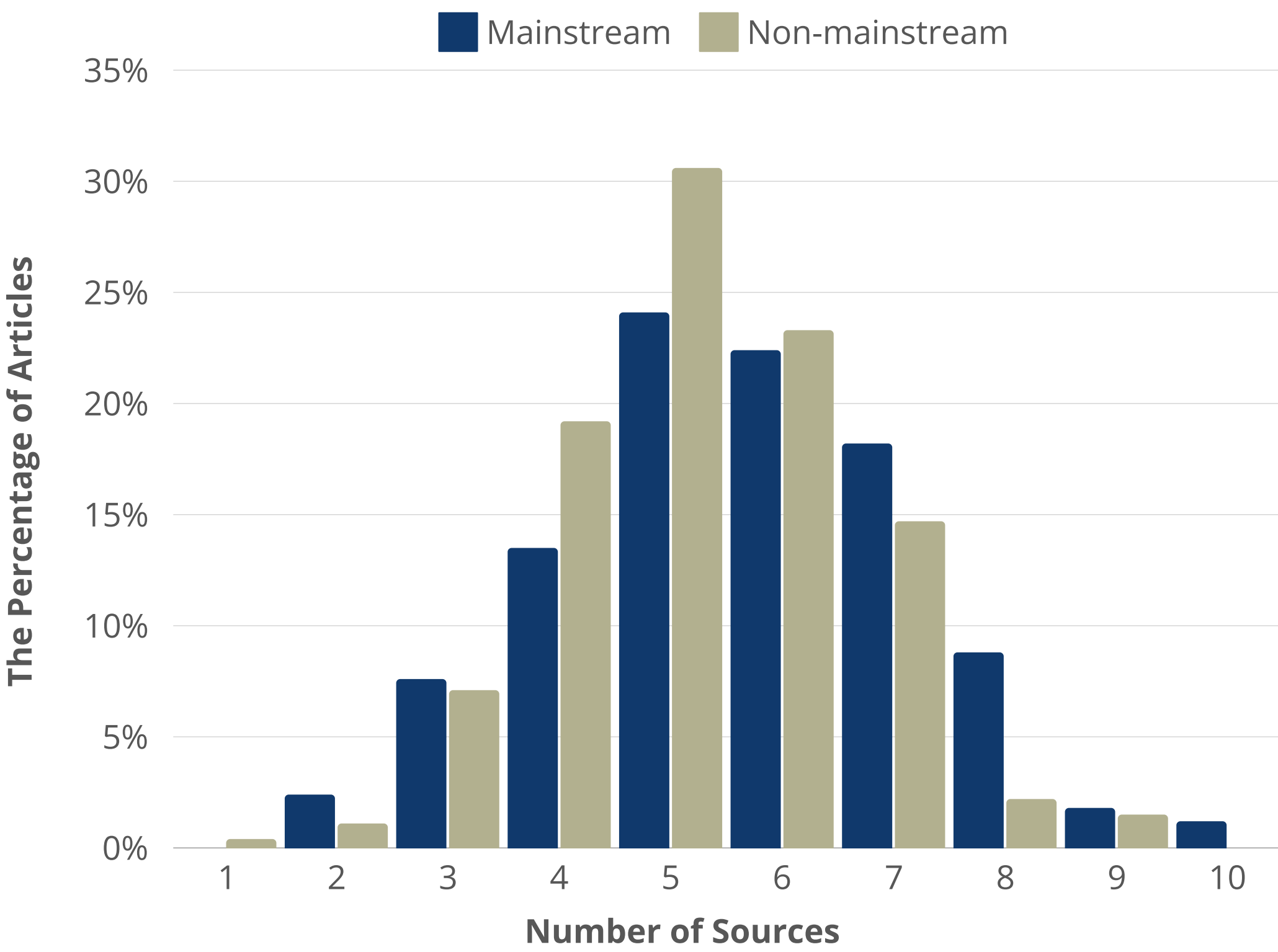
DATA ANALYSIS



The Times of Sources Being Reported

There were 17 types of sources in the articles. **“Other organization”** was the source reported the most among all articles. We can focus on seeing that **“Victims”** was the second large sources being reported. Also we notice that article took many perspectives from **“Politicians”** and **“Professionals”**.

DATA ANALYSIS



The Distribution of Sources Count

Typically, most of the articles used 5 sources. Over **90%** articles used **more than 3 sources**, which meant that the articles were trying to make their content more reliable with different opinions. The distributions of number of sources in the mainstream and non-mainstream were similar. Both of them had a bell-shape distribution. Number of sources used in **mainstream was more various**.

CONCLUSION

Target Audience

- Qualitative researchers
- Any type of writers
- Government departments
- Those who need public opinion



Features

- Summary or articles
- Coding themes & sources with AI
- Visualization of figures and tables



System Benefit

- Saving time and labor cost
- Objective perspectives
- Providing novel ideas



THANK YOU

● FOR YOUR NICE ATTENTION



December 2024