

Article

Seeing Bias at a Glance: A Visual–Statistical Analysis of Sentiment in China’s State-Backed English News Media

Xiangning Liang

Independent Researcher, Nottingham NG7 3LP, UK; xiangning.liang@outlook.com

Abstract

China’s state-backed media is valuable for news bias research due to the tight control of journalism in China. In the digital era, bias remains, and quantitative and computational methods are playing an important role in studying it. Bias on China’s English news websites has not been examined in previous research, and a day-to-day angle is lacking. This study selects four well-known news media websites in China: CGTN, China Daily, Global Times, and Xinhuanet, which are owned and operated by the state or party in different ways. The BBC is chosen as a benchmark of editorial independence to highlight differences in bias. The news titles on their official websites were collected on a daily basis and analysed with sentiment as the focus. Features of news webpages are discussed and utilised. The charts and network graphs in this paper lower the barrier to comprehension for wider audiences, enabling readers to grasp the sentiment bias of news media in a visually digestible format. The results demonstrate that sentiment bias exists in China’s state-backed English-language news websites today, favouring positive coverage of the domestic side. In contrast, the BBC serves as a suitable benchmark and reflects the tendency for negativity dominance in news reporting.

Keywords: sentiment bias; digital journalism; news collection; data visualisation; China; state-backed media; English news websites; news titles



Academic Editors: Yao Song and
Andreu Casero-Ripollés

Received: 15 June 2025

Revised: 9 August 2025

Accepted: 7 October 2025

Published: 15 October 2025

Citation: Liang, X. (2025). Seeing Bias at a Glance: A Visual–Statistical Analysis of Sentiment in China’s State-Backed English News Media. *Journalism and Media*, 6(4), 177. <https://doi.org/10.3390/journalmedia6040177>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Chinese authority’s ruling power is exerted in many domains, for example, elections, law enforcement, judiciary, media, economy, civic space, etc. News media is one of the most accessible ways for researchers and the public outside of China to know what is happening in this country; it is also a domain in which the authority imposes control over what can be exposed and what the public can see. In an era where journalism has gone digital, evidence of such control can still be found, and bias is one of its signals. This research chooses English news websites of four representative state-backed media outlets in China as a focus and the BBC as a contrast; this selection provides an external benchmark that helps wider audiences understand how the selected Chinese news media outlets differ from a world-renowned and respected one. The method used and described in this paper is adaptable for long-term news tracking purposes. The project dataset is shared. Plus, the body of this paper is essentially math-free and image-rich to show the power of visualisation. This work not only presents media bias in an intuitive way but also demonstrates that daily news data is effective in bias identification. Future news research is expected to combine both time granularity and topical categories, inspired by the results and limitations presented in this paper. The following sections review relevant

studies about media bias related to China, theories behind bias, and research motivation and objectives.

1.1. Media Bias Related to China

Media bias around China is an enduring topic in news and media research, and many patterns have been discovered. In 1988, [Warren \(1988\)](#) studied foreign and domestic news content by analysing stories from China's newscasts. The results reveal that success in economics and industry makes up the biggest proportion of domestic stories, while international relations and crime events dominate international stories. Warren attributed the findings to the centralist reality of the Chinese system.

Using content from various newspapers from the local level to the central government level, [Qin et al. \(2014, 2018\)](#) concluded that newspapers with closer institutional ties to the central government tend to have more propaganda and political content. Bias can also be found by comparing foreign news on television in mainland China, Hong Kong, and Taiwan. Top leaders' speeches and the country's world influence carry greater weight on mainland China's television, but Hong Kong and Taiwan have more reports on internal politics and social groups ([Lin et al., 2011](#)).

Sentiment bias is a relatively new analytical angle in news media research. Such bias has been used to add barriers to foreign automakers in China to protect the government's interests. [Kim \(2018\)](#) used a sentiment dictionary to calculate sentiment scores for car-recall reports in newspapers, and she found that the reports on foreign automakers are at a more negative level. China's official newspapers maintain such bias to influence consumers' impressions of non-state-owned automotive enterprises. [Schweizer et al. \(2025\)](#) tested the relationship between the political connections of local firms and media bias with machine learning-based sentiment analysis. They identified that, for those firms with a tight political bond with local government, the media tends to build positive images while suppressing negative news about them. Yet, politically unconnected companies do not have such an advantage.

The effects of news bias were examined by [Shen and Guo \(2013\)](#) based on a survey, which proved that the sentimental positivity of news has a positive correlation with the audience's political trust and national pride. The success of bias can be considered a natural result of Chinese people's perception of the high credibility of state-controlled sources, though such credibility does not match people's actual readership of those media ([Zhang et al., 2014](#)). More interestingly, in another survey experiment, respondents showed a clear awareness of the existence of media bias, but they still expressed their preference for official media and chose not to trust foreign outlets ([Truex, 2016](#)).

When studying media bias, dynamics is another important aspect worth considering, since the media environment in China has changed during the past decades. Conglomeration reform is an example. Fetters on some nonofficial media were released, and commercial media emerged. [Piotroski et al. \(2017\)](#) pointed out that the positive tone increased in official articles and decreased in nonofficial ones during the reform. Nonofficial media started to have more freedom in reporting market-oriented information, though this does not mean that they can criticise the policies of the central government or make comments on international affairs freely. Moreover, this progression might not be able to continue. Some scholars argued that the fruit of conglomeration reform started to be destroyed by Xi's tighter control. In terms of reasons, [Singh \(2016\)](#) highlighted the Ren Zhiqiang case to show the fears of the regime toward liberalised arguments by celebrities in the media.

Social media has more interactivity compared with television and newspapers. Utilising social platforms, China's official channels implement the "Going-out" policy by spreading positive discourse and affecting audience engagement with the trick of senti-

ment bias. On Facebook, China's accounts applied different tones to BRI (the Belt and Road Initiative) countries with higher positive sentiment levels. [Liang \(2019\)](#) argued that this strategy's purpose is to use these news-like posts to show "China's role as a builder of world peace" and as a helper to BRI countries. This finding is supported by the results of applying data mining to posts, likes, and comments under the official accounts on Facebook.

Many types of biases are more obvious and easier to spot due to the Great Firewall and national control of news outlets. Here are more patterns found in research based on Twitter (now X). [Huang and Wang \(2019\)](#) said that the major source of diplomatic posts is Chinese official media, which produced more than twice as many as its foreign counterparts, reflecting Xi's "Tell China Stories Well" concept. Central-level state accounts dominate among actors (accounts that reply, mention, or like a topic) around the relevant issues, and those accounts rarely interact with other users, as [Nip and Sun \(2022\)](#) concluded, after analysing 154,542 tweets with the #SouthChinaSea tag. An Oxford Internet Institute study ([Schliebs et al., 2021](#)) identified a coordinated account network on Twitter that amplified the discourses of China's diplomats. Many accounts were fake or inauthentic. The data shows that this network contributed 44% of retweets and 20% of replies to ambassadors' posts, which is a covert pattern that casts doubt on the truthfulness of the image of China's diplomatic outlet on Twitter. From these studies, it can be learnt that, in addition to imposing direct control and large-scale censorship at home, which have been widely identified ([Xu & Albert, 2014](#)), China has been trying to build its positive image in the world arena, where other participants are less likely to compromise under its power.

1.2. Theories Behind Bias

Summarising the aforementioned research, some implicit rules in China's media seem to emerge from the evidence. In an environment that is covered by China's administrative forces, China's image is improved, and foreign countries' images are degraded. On foreign social platforms, such force is applied to China's accounts or channels. The difference is that the Chinese regime is the "big brother" in the domestic space, where the local media and foreign media are pawns in its hands, while in the international arena, China is just one of the players; it exerts power to strengthen its own side. [He \(2008\)](#) described media control in China as "putting out more positive reports and fewer or no negative reports". [Stockmann and Gallagher \(2011\)](#) adopted the term "positive propaganda" to summarise the media report patterns. [Brady \(2006\)](#) borrowed Deng Xiaoping's dictum "stress stability above all else" to explain such propaganda as a natural continuation of Chinese-style socialism theory. Media framing is another angle to explain. [Sheafer and Dvir-Gvirsman \(2010\)](#) discussed the different effects of negative and positive framing. Negative framing can make a much more significant impact on public reaction than positive framing, and it may attenuate the government's power to rally the public in support of a peace process. A psychophysiological experiment also proved that negative news elicits a more intensive reaction than positive news by measuring participants' physical parameters from their bodies during news reading ([Soroka & McAdams, 2015](#)). This theory can explain China's media bias in another way. Firstly, audiences are expected to hold a positive and stable image of China, which is why positive news is highlighted, and negative news is suppressed. Secondly, readers are expected to have a turbulent or controversial image of the world outside of China. The regime does not mind, or it acquiesces to the public's more emotional opinion toward other countries. It can be said that this strategy aims to maintain a persistent sentimental difference in the media so that it can eventually affect the opinion of the general public.

By way of analogy, China is playing a game of tilted balance with the media, where the goal is to maintain its vantage on the "right" side in different scenarios. Between central

and local sides, the central side is the “right” side. Between domestic and foreign sides, the domestic side is the “right” side. To keep such a vantage, more positive language is used, and more positive events are reported. In addition to explaining this with the term “strategy”, it can also be explained by the habits of communist authorities, as they often glorify their own images and undermine the images of hostile countries, especially in periods when the ideology within the communist party was extremely left, like the Soviet Union in the Cold War and China in Mao’s era.

1.3. Motivation

Even though media bias has been widely explored, day-to-day analysis of news coverage is still lacking. Warren (1988) suggested that a longitudinal comparison of China’s foreign news coverage with coverage from another country might be valuable for scholars seeking to understand differences. However, researchers have not put much effort into this approach. Media articles were organised by year in some projects (Morstatter et al., 2018; Dai & Luqiu, 2022), but this is still a large timespan. Continuous and granular media data is an angle that has received insufficient attention from researchers. The reason for this may be because of the dataset. Records of television and newspapers are not usually formatted or digitalised, and precise timestamps are frequently unavailable. Therefore, the most efficient way to spot patterns is to treat the data as a whole corpus and categorise it into different topics, such as politics, culture, and business. Apart from topical categorisation, some bias research focuses on comparing event-driven reports from China and other countries like Africa during COVID-19 (compared with Western countries) (Gabore, 2020), vaccination (compared with the US) (Ju et al., 2023), and Koizumi’s visits to Yasukuni Shrine (compared with Japan) (Tan & Zhen, 2009). These studies are issue-specific research (Dekavalla, 2018) because only event-relevant discourses are selected. Through a topical or issue-specific lens, patterns of individual media cannot be readily compared, while such comparisons are important if researchers want to understand media within a more comprehensive scope.

To understand bias in China’s news media and to know how the pattern is different from an external reference media source from a general angle, as well as to attempt to organise data by day, news websites were considered appropriate targets. Firstly, research on bias on China’s news websites remains underexplored, and webpage-based patterns remain unknown. Moreover, news websites serve as one of the most direct channels to study individual media outlets. Secondly, news websites are suitable for data collection. Automated web scraping scripts can handle such tasks. Scheduled script execution can fetch the content that is displayed to viewers. Thirdly, information on news websites has already been categorised, which means that news reports are usually placed under a certain column or section; this common feature can be used to analyse differences in bias patterns across multiple news websites. Sentiment was chosen as a measuring method because it can be applied to a large dataset and reflect news positivity and negativity efficiently. There are well-developed packages developed for sentiment analysis, thanks to advances in natural language processing by computer scientists and linguists.

Another important point is data visualisation. Most previous research used tables, numbers, formulas, and text examples to demonstrate their findings, but they failed to convey a straightforward insight to wider audiences like journalists and new students of journalism. Too much time needs to be spent on interpreting complex evidence when readers are trying to understand the arguments and patterns. Visualisation is a sharp tool for showing findings in social science and for storytelling: it offers vivid and intuitive ways to convey information, such as line charts, dot charts, multiple bar charts, network graphs, etc. In journalism and media research, many previous studies have demonstrated

the effectiveness of visualisation methods (Schliebs et al., 2021; Dimitriadis et al., 2024; Witzenberger & Pfeffer, 2024; Samalis et al., 2023).

1.4. Research Objectives

Building on the preceding discussion of relevant studies, research gaps have been identified. News bias around China in traditional and social media has been well documented, but bias on news websites as an intermediate platform is relatively unexamined. Evidence of news bias was mainly discovered from topical and issue-specific research using data gathered from a large timeframe. Exploring news bias at a day-to-day level from individual media introduces a new analytical angle. Together, inspired by the power of web data collection, sentiment analysis, and visualisation, this study aims to explore the following questions: (1) What are the sentiment patterns in China's state-backed news websites today when reporting on domestic and international domains, as demonstrated on a day-to-day basis and supported by visualisation? (2) Can sentiment bias be confirmed using data from China's state-backed media itself, as well as through comparison with a benchmark media organisation?

2. Materials and Methods

This study adopted a dichotomous methodology to investigate the sentiment patterns of selected official news websites in their domestic and international coverage. The following sections discuss the selected news media, where and how the data were collected, why the news title was chosen, and data analysis, as well as a description of the dataset.

2.1. Data Collection and Scope

Four of China's English-language news websites were selected: CGTN—China Global Television Network (the international division of China Central Television); China Daily (owned by the Publicity Department of the Chinese Communist Party); Global Times (under People's Daily); and Xinhua (state-owned, under the State Council). Although China Daily and Global Times are officially affiliated with the Chinese Communist Party, their operations are deeply integrated with state-owned companies. In this study, they are treated as state-backed media.

The choice of benchmark media is intended to reveal a distinct sentiment pattern that can help to highlight the potential bias in China's state-backed English media. Five criteria were considered when selecting the benchmark: website layout, language alignment, governmental involvement, global engagement, and editorial independence. Among these, editorial independence serves as the key variable, while the remaining four establish common ground with the selected Chinese state-backed media. The BBC News website is the ideal option. It is a statutory corporation with a national-level mandate, and it is known for editorial independence worldwide. It also shares certain website layout features with the selected Chinese media. As Figure 1 shows, the layouts of their webpages are highly similar, with a homepage and several columns that contain news reports about specific topics. "China/UK" and "World" columns can be found on each website, as they host domestic news and foreign news, respectively. To collect news under the columns every day, web scraping scripts were executed once per day (see Table 1 for URLs).

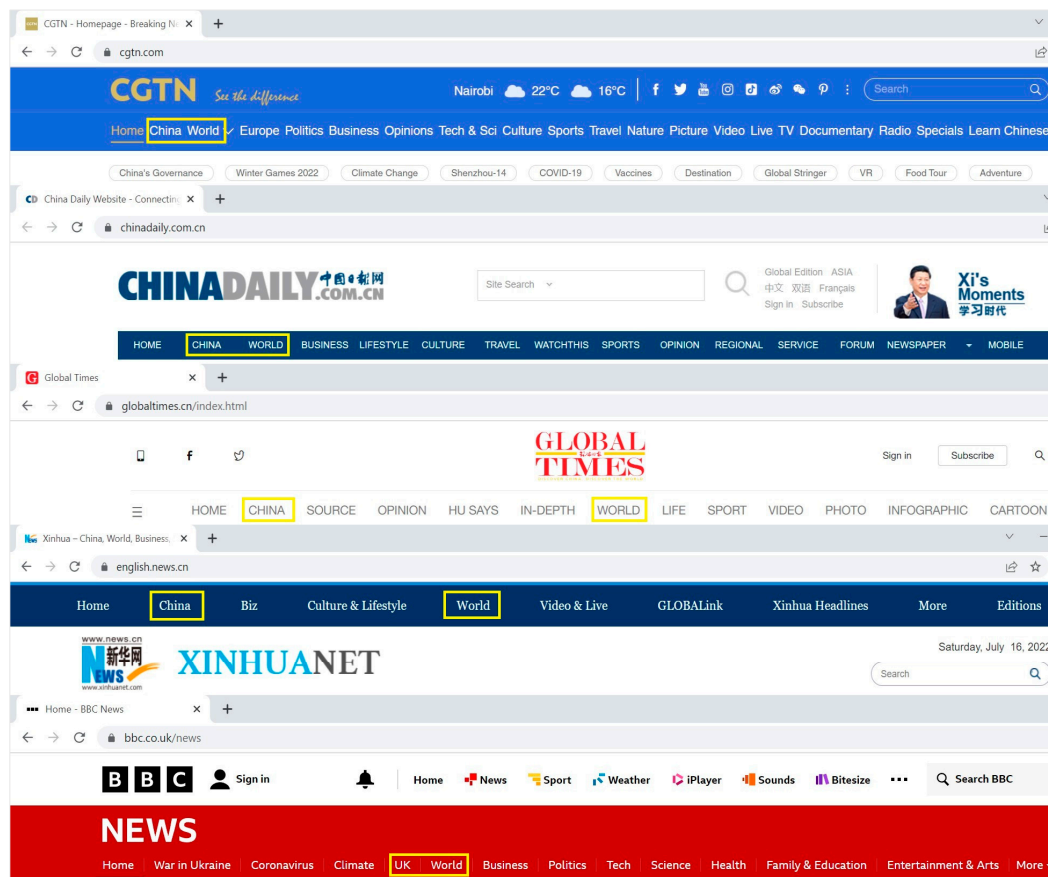


Figure 1. Homepages of selected news websites. Domestic and world columns are highlighted.

Table 1. Selected media and news website links of domestic and world columns.

Media	Domestic Column	World Column
CGTN	www.cgtn.com/china	www.cgtn.com/world
China Daily	www.chinadaily.com.cn/china	www.chinadaily.com.cn/world
Global Times	www.globaltimes.cn/china/index.html	www.globaltimes.cn/world/index.html
Xinhuanet	https://english.news.cn/china/index.htm	https://english.news.cn/world/index.htm
BBC	www.bbc.co.uk/news/uk	www.bbc.co.uk/news/world

Note: All websites accessed on 8 September 2022.

The news title is the content of interest. Under the column, the news title is one of the most eye-catching elements on listing pages that aggregate multiple news items (see Figure 2). This is a shared feature among the five websites. The news title was chosen as the research object for the following reasons. The news title is an essential device in media framing (Tankard, 2001). Researchers found patterns and reflected social topics by analysing news titles (Liu et al., 2019). Psychologically, the title plays a significant role in readers' behaviours in online news consumption; it affects readers' first impression and cognition in many aspects (Montejo & Adriano, 2018; Lagun & Lalmas, 2016; Condit et al., 2001; Kuiken et al., 2017; Rieis et al., 2015). Based on these studies, analysing news titles is expected to reveal bias, if present.

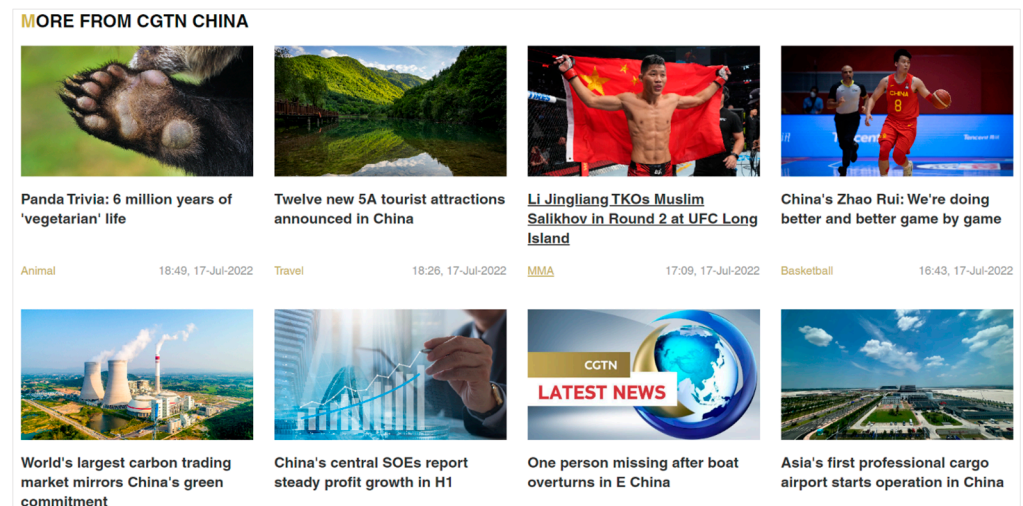


Figure 2. Layout of news webpage (screenshot from CGTN). Visual predominance of news titles.

There are many elements of a news webpage apart from news titles and images, such as banners, advertisements, social media links, and other hyperlinks that lead to external websites. By analysing web source codes, HTML tags that contain news title content can be found. News titles can then be located and extracted after parsing the retrieved webpages (see Figure 3).

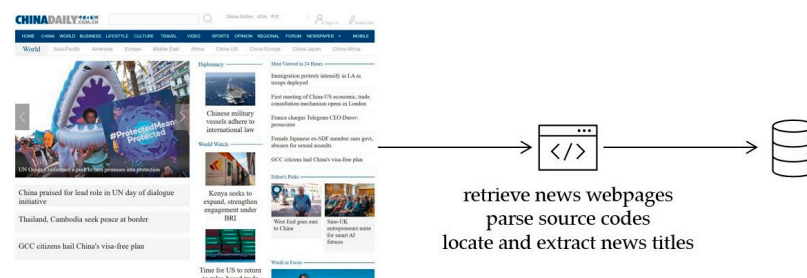


Figure 3. News title collection. From webpages to local storage through scripts.

2.2. Dataset Description

Collected data are stored in CSV file format. In addition to the title of a news report, the rest of the logged attributes are the publishing time of a news title, the column where a news title resides, and the date when a news title is collected. Compared with traditional news media (newspaper, television, radio), a news website has unique features. Traditional news media has almost no repetitive reports from day to day: even for long-term or follow-up coverage of the same event, it is hard to find two identical reports. Conversely, for news websites, a single news report and its title can persist on the webpage for many days. The news report's lifecycle depends on the website's operation. Table 2 shows more information about the dataset. The complete dataset can be downloaded from Supplementary Material (S1).

Table 2. Dataset description.

Media	Time Range (2022)	Num of Entries	Unique Entries
CGTN	19 February–8 September	69,805	14,814
China Daily	19 February–8 September	13,984	4559
Global Times	19 February–8 September	42,361	5893
Xinhuanet	19 February–8 September	18,553	10,511
BBC	30 June–8 September	4286	1324

2.3. Data Analysis

Data analysis was conducted with Python 3.1.0 and the Vader 3.3.2 sentiment analysis package (Hutto & Gilbert, 2014). The sentiment analysis function in this study can be briefly explained in this way: text (news title or single word) in, sentiment score out. The sentiment score ranges from -1.00 to $+1.00$. The magnitude indicates the high/low strength of sentiment, and the sign indicates positivity/negativity. A sentiment score of zero indicates that the input text is considered neutral. During data analysis, the same pipeline was applied to the data from every selected media outlet to ensure the validity of the comparison. The sentiment analysis was conducted in four dimensions: sentiment in total news output, sentiment in everyday news page, sentiment in daily news production, and sentiment in vocabulary use. The sentiment in vocabulary use was analysed token-wise; the other three dimensions were title-wise.

3. Results

In this part, findings are illustrated with visualisations and descriptions, along with explanations about how data were handled. Sentiment patterns of China's state-backed English media are first presented, followed by those of the BBC and a summary.

3.1. Sentiment in Total News Output

This section characterises the sentiment pattern based on all news titles collected from individual media websites during the project and grouped column-wise. Duplicate titles have been removed. Data points with a sentiment score of zero or neutral titles are excluded. Figure 4 shows the distribution of sentiment in the domestic and world columns of the four Chinese news websites. Every figure consists of a swarm plot and a violin plot. White dots are data points of sentiment scores. Violin plots draw the outlines of swarm plots, which can reflect the distribution in a more intuitive way. The length of the lines of white dots and the width of the violin shape represent the relative number of data points at a certain sentiment level.

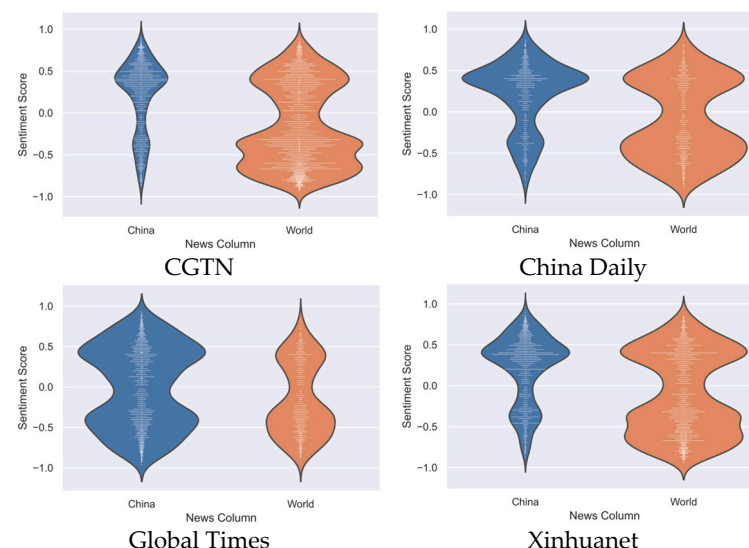


Figure 4. Sentiment distribution of total news output.

It can be seen that CGTN, China Daily, and Xinhuanet tend to output significantly more positive news than negative news in the domestic column, marked by hammer-like shapes with big heads and thin handles. Global Times shows a more balanced pattern when reporting domestic news. For the world column, the four news sites output more negative than positive news, marked by gourd-like shapes.

3.2. Sentiment in Everyday News Page

This part of the results is about day-to-day sentiment levels. The web scraping scripts were run once a day to collect news titles, which are used to represent the sentiment level of that day. Column-wise grouping is applied.

Figures 5–8 show the scatter plots of sentiment. Scores of news titles collected on the same day are drawn on the same vertical line. Data points with bigger absolute values are coloured darker (neutral data points are transparent). The blue line connects the average sentiment values of each day. The orange line connects the average sentiment values of each day without neutral data points. The density shown in the figures reflects the distribution of sentiment.

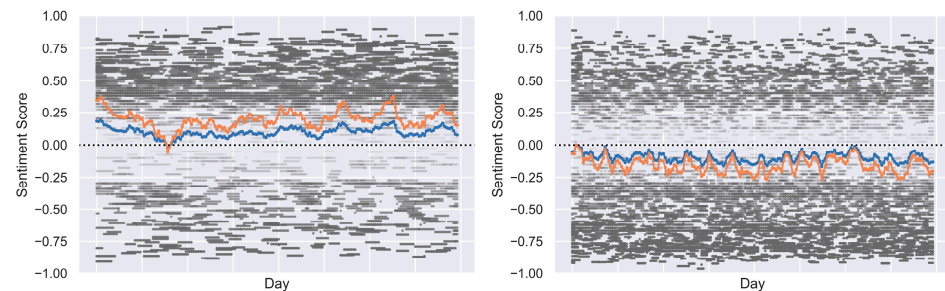


Figure 5. Daily sentiment level on page. CGTN. Column: **(Left)** China, **(Right)** World. Blue: average with neutral data; orange: average without neutral data.

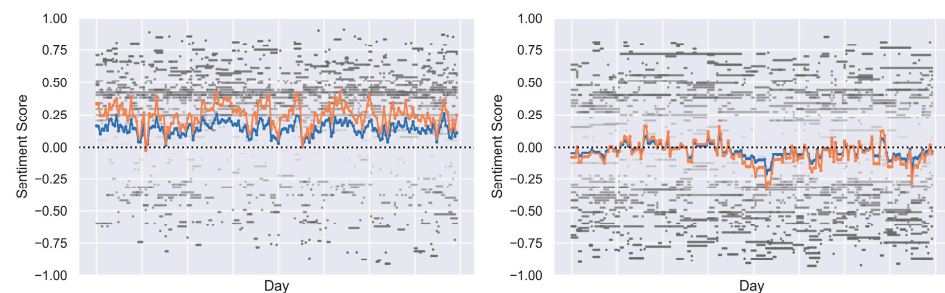


Figure 6. Daily sentiment level on page. China Daily. Column: **(Left)** China, **(Right)** World.

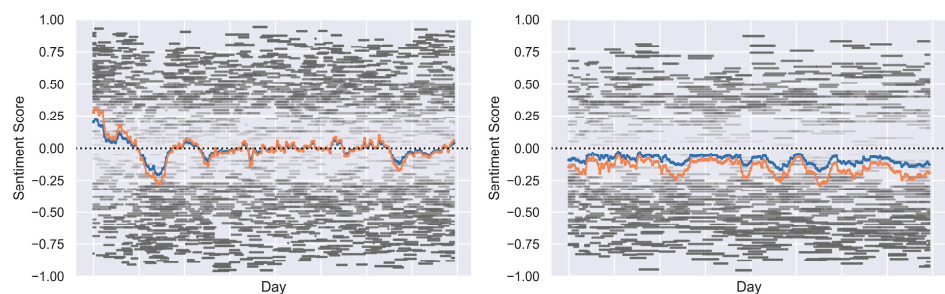


Figure 7. Daily sentiment level on page. Global Times. Column: **(Left)** China, **(Right)** World.

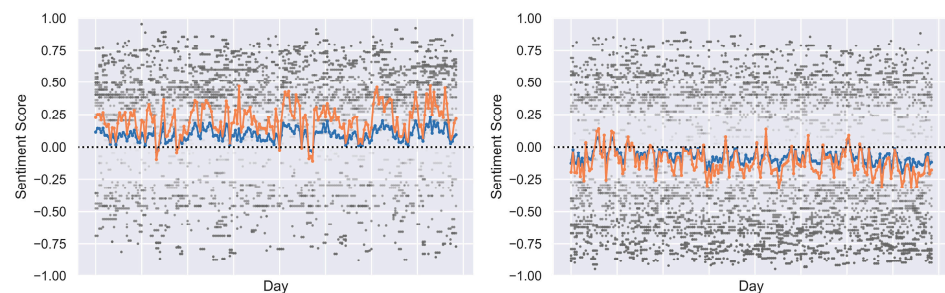


Figure 8. Daily sentiment level on page. Xinhuanet. Column: **(Left)** China, **(Right)** World.

The results illustrate that the CGTN and Xinhuanet websites have a similar pattern in everyday news sentiment presentation: the average sentiment value of the China column is always positive, while that of the world column is always below the zero line. China Daily and Global Times do not follow this pattern completely, yet they each reflect a different half of the pattern. China Daily has a clear preference that puts more positive news over negative news in the domestic column, while the average lines of the world column cross the x-axis several times. Global Times has a clear preference that puts more negative news over positive news in the international column, while the average lines of the domestic column cross the x-axis many times with significant oscillation.

3.3. Sentiment in Daily News Production

This part describes the sentiment pattern associated with the news publishing day. A publishing time stamp under the title is a shared feature among the selected websites (see Figure 9). As discussed previously, news can persist on webpages for more than one day. The previous section focuses on the sentiment of news titles displayed on the webpage, and this section groups them by their actual publishing day.



Figure 9. Posted news title and publishing time stamp (screenshot from Global Times).

Stacked histograms are used to show results (see Figures 10–13). Each bar in the graph represents news produced on the same day. To show proportions, the bar is filled to the top. The two lines in each graph are kernel density (KDE) lines that show smoothed distributions of the lower two categories. The KDE lines and colours separate the graphs into three areas: negative, neutral, and positive.

Except for Global Times, the other three news websites have a significantly bigger proportion of positive coverage than negative coverage in the domestic column. For the world column, the pattern is the opposite, including for Global Times. For the domestic column of Global Times, every day, the proportions of positive and negative news production are more balanced.

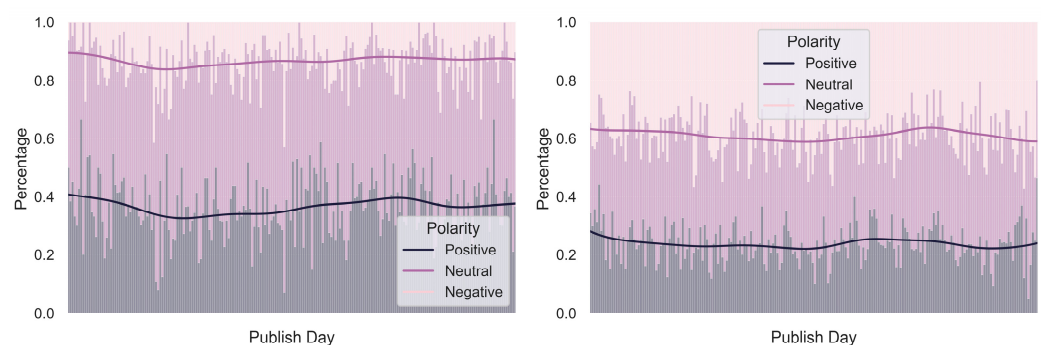


Figure 10. Sentiment of daily production. CGTN. Column: (Left) China, (Right) World.

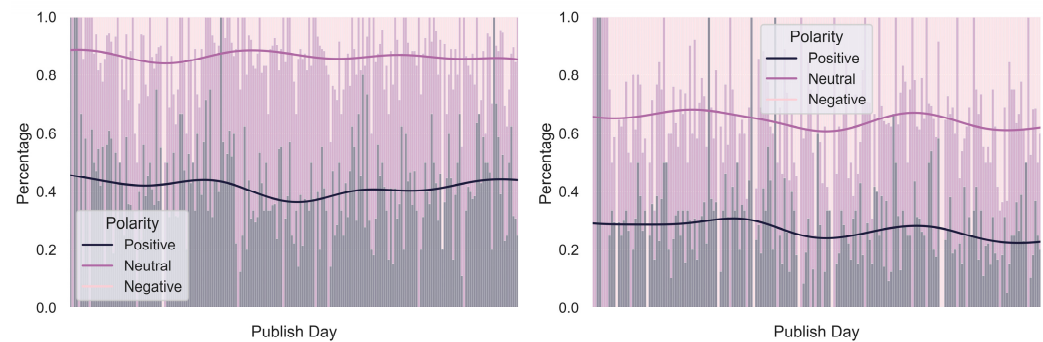


Figure 11. Sentiment of daily production. China Daily. Column: (Left) China, (Right) World.

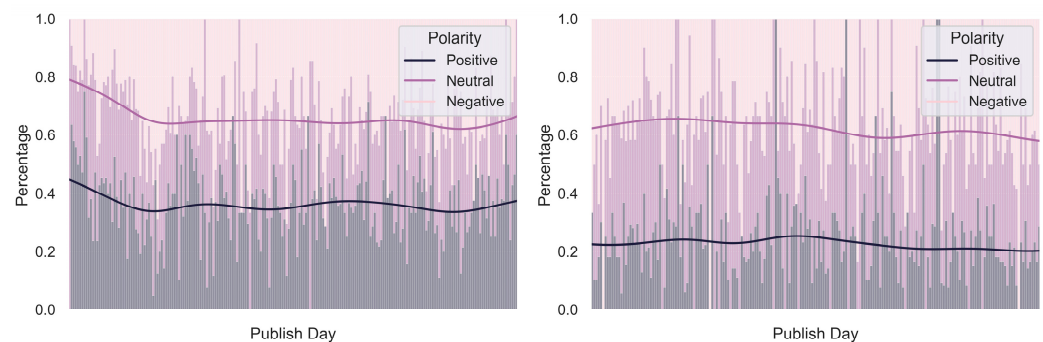


Figure 12. Sentiment of daily production. Global Times. Column: (Left) China, (Right) World.

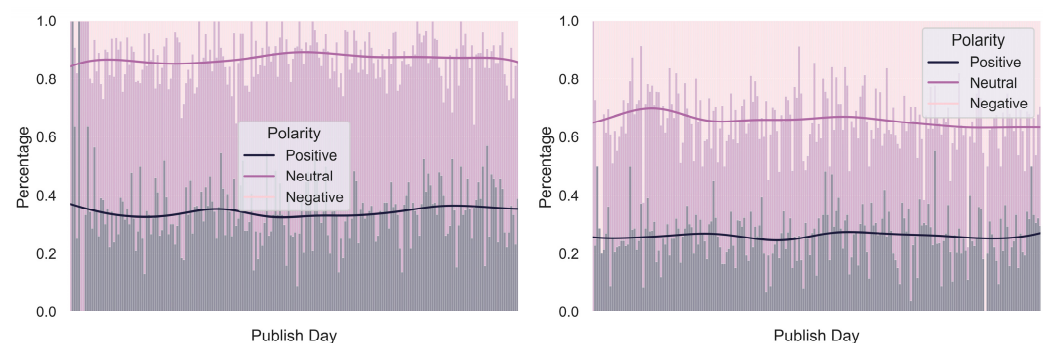


Figure 13. Sentiment of daily production. Xinhuanet. Column: (Left) China, (Right) World.

3.4. Sentiment in Vocabulary Use

This section concentrates on news titles that explicitly contain names of countries or nationalities. If a title includes the specific name of a country or nationality, it is more geopolitical than those that do not. The reason is that such a title-writing style can emphasise an event's national factor or highlight one's national background. With this insight, more precise filtering and classification can be performed, without relying on the website's domestic and world column categorisation. Two groups of news titles remain after filtering and classification. Group 1: Titles mentioning China/Chinese; Group 2: Titles mentioning other countries/nationalities. Each title in both groups contains only one mention of a country or nationality. Titles that contain multiple country or nationality references are excluded to ensure clarity and reduce confounding sentiment influences. This was achieved by using an exhaustive list of geopolitical tags that include names of countries and nationalities, as well as common abbreviations (e.g., United Kingdom, UK; United States, US, American; China, Chinese; Germany, German).

The sentiment of words associated with a country or nationality is visualised using network graphs. Titles are broken down into tokens, and stop words are filtered out. The sentiment tool assigns each token a score; neutral words are excluded. Tokens are connected with a country/nationality that appears together with its title. The country/nationality is encoded as the source, and the token is encoded as the destination. Positive tokens are coloured in green, and negative tokens are coloured in yellow. Nodes of China and Chinese are coloured in red; other countries/nationalities are coloured in blue. Lines connecting source and destination nodes are marked by the colour of the source node. The size of a token node depends on the absolute value of its sentiment score. The word (green and yellow) nodes in networks have been configured to be clustered and gravitate around the countries/nationalities (red and blue) nodes it is connected with.

Figures 14–17 show the network graph for each media site. Also, see Figures A1 and A2 (in Appendix A) for a closer look. In all four networks, green node clusters can be easily found around “China” and “Chinese” nodes. It indicates that positive lexicons are largely assigned to news titles in which China/Chinese is explicitly mentioned. Yellow nodes scatter around other country/nationality nodes.

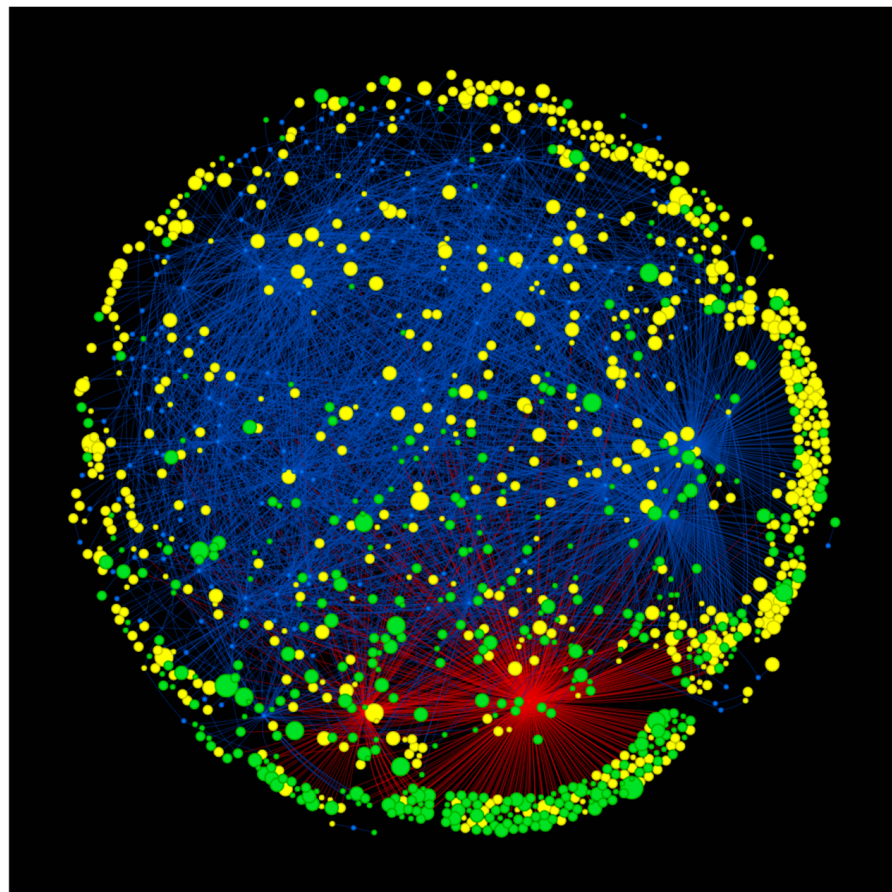


Figure 14. Complete vocabulary sentiment network. CGTN. Green and yellow nodes represent vocabulary tokens, with size proportional to sentiment magnitude (green = positive, yellow = negative). Red and blue nodes represent countries/nationalities, with identical sizes (red = domestic, blue = foreign).

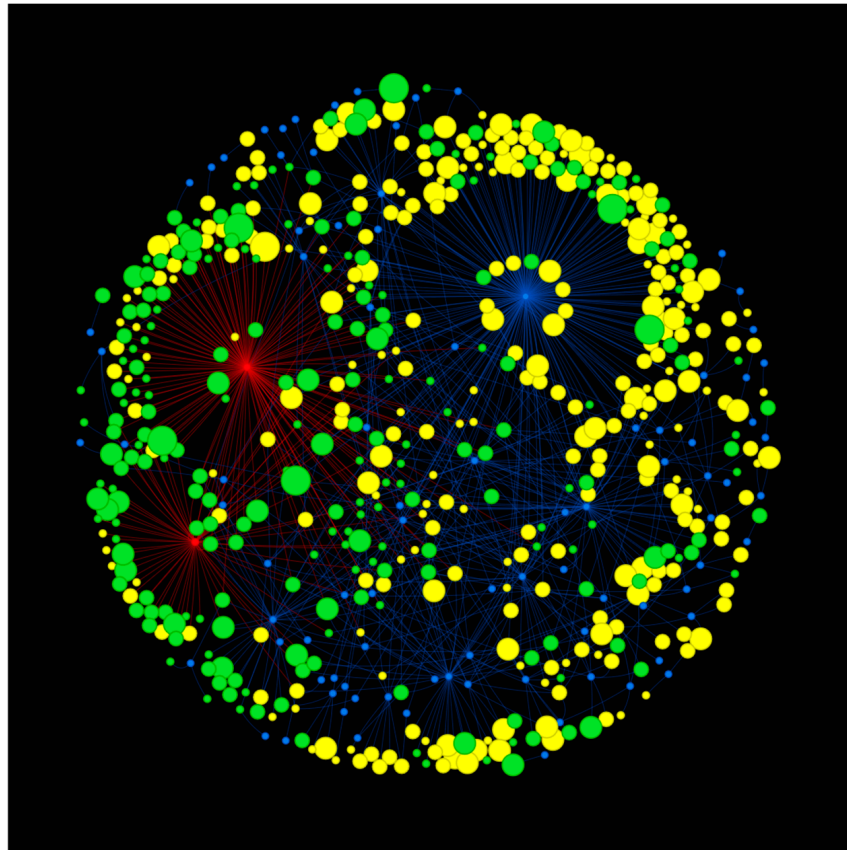


Figure 15. Complete vocabulary sentiment network. China Daily.

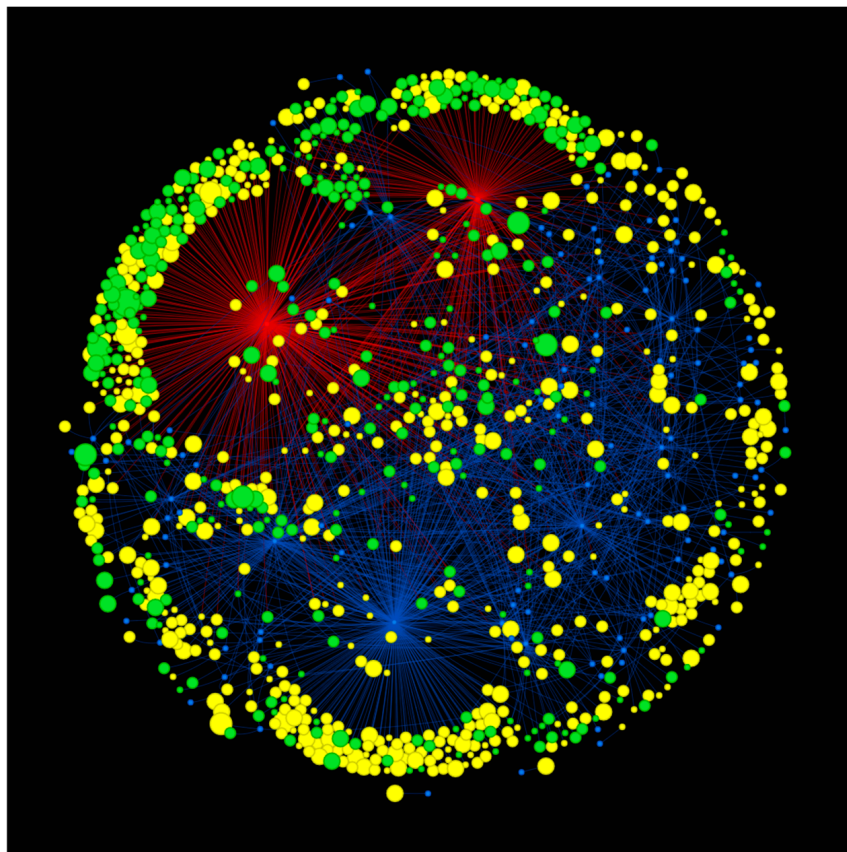


Figure 16. Complete vocabulary sentiment network. Global Times.

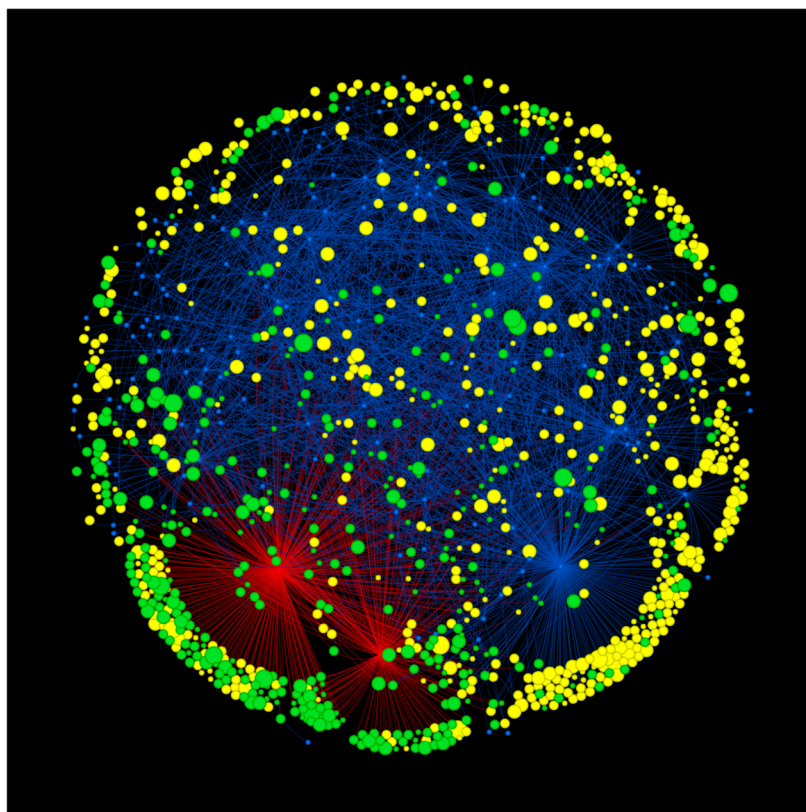


Figure 17. Complete vocabulary sentiment network. Xinhuanet.

3.5. Sentiment Results for BBC

Figures 18–21 show the visualisations based on BBC data. In Figure 21, the nodes for “UK,” “United Kingdom,” “British,” “Scotland,” “Scottish,” “Wales,” “Welsh,” “England,” “English,” “Northern Ireland,” and “Northern Irish” are marked in red, reflecting the UK’s nature as a union state composed of distinct national identities. The BBC did not demonstrate a similar sentimental pattern to the four Chinese news websites. It is basically a pattern of negative dominance, where positive elements make up only a small proportion, in the coverage of both domestic and world scopes. Even a balanced pattern is absent. Unlike China’s news pages, which prefer positive titles for domestic news, the BBC’s news tends to use negative ones in both columns.

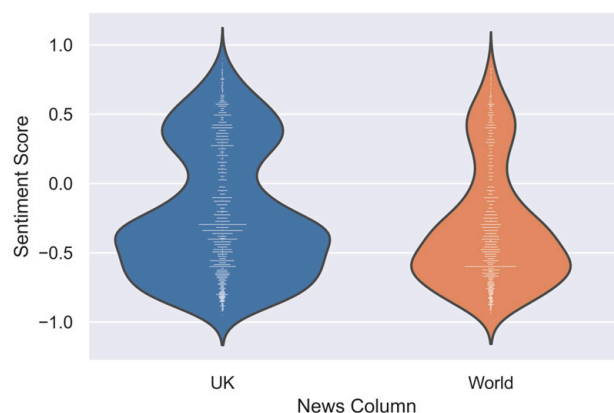


Figure 18. Sentiment distribution of total news output. BBC.

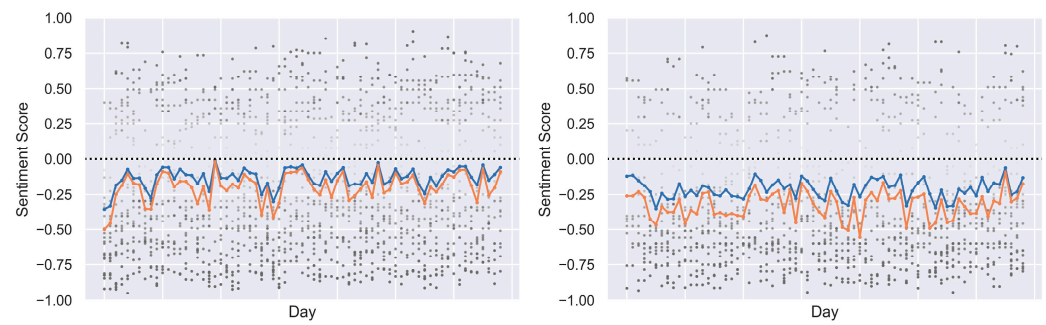


Figure 19. Daily sentiment level on page. BBC. Column: (Left) UK, (Right) World.

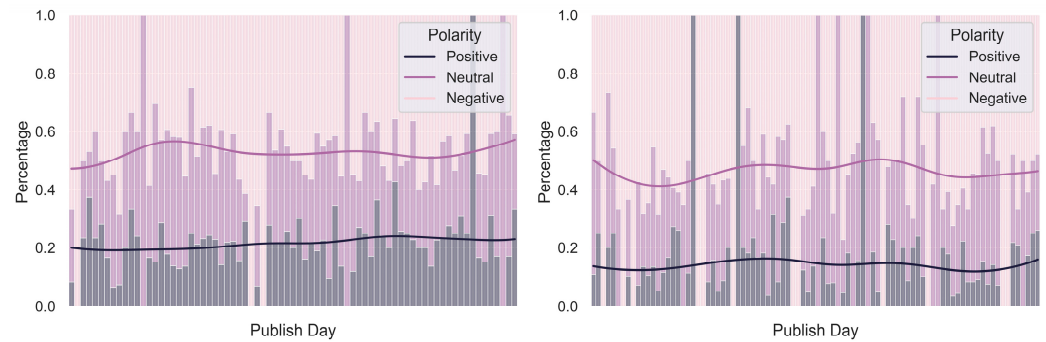


Figure 20. Sentiment of daily production. BBC. Column: (Left) UK, (Right) World.

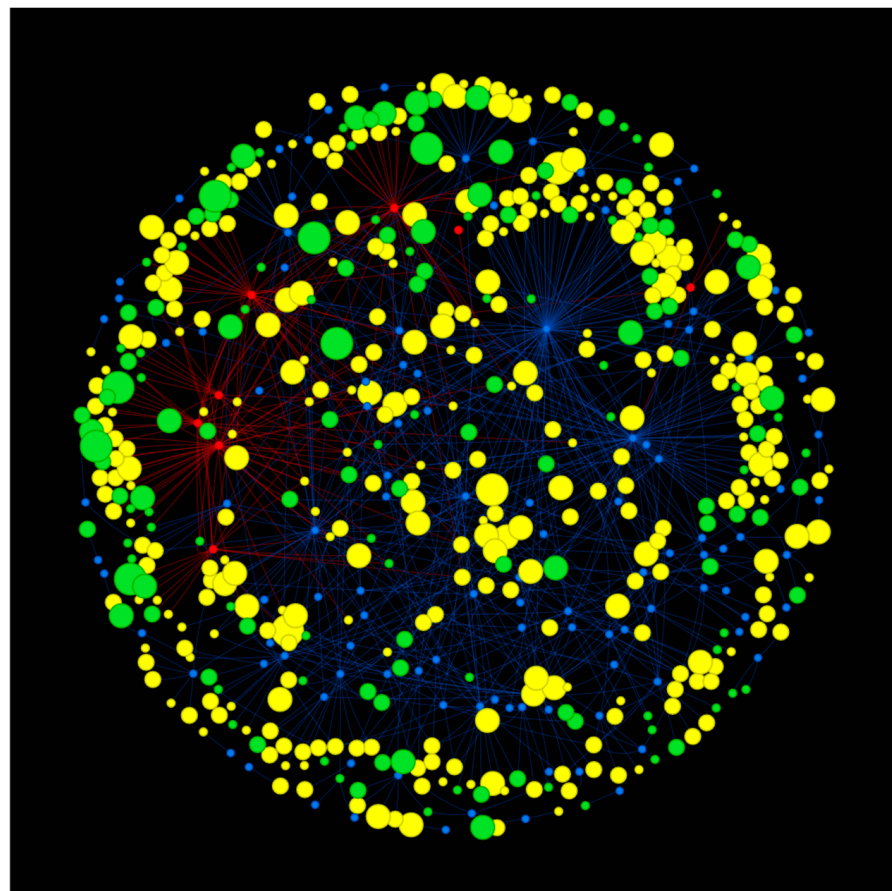


Figure 21. Complete vocabulary sentiment network. BBC.

3.6. Summary of Sentiment in China's State-Backed Media and BBC

The visualised findings are summarised in Table 3. A simple labelling scheme is used to represent the sentimental tendencies discussed in each section. If a graph illustrates a clear sentiment tendency in a news column, it is marked as positive + or negative −. If a graph does not show an obvious tendency, the column is marked as balanced or neutral with ○. The table also points out whether a news website has a “Positive Inside (domestic), Negative Outside (foreign)” tendency, which means a typical pattern of building a positive image for the home country and a negative image for the rest of the world in general. Sentiment patterns and findings are further discussed in the next section.

Table 3. Summary of sentiment patterns from previous sections. +: dominated by positivity. −: dominated by negativity. ○: relatively balanced or neutral.

		CGTN		China Daily		Global Times		Xinhuanet		BBC	
		China	World	China	World	China	World	China	World	UK	World
3.1.	Sentiment in Total News Output	+	−	+	−	○	−	+	−	−	−
3.2.	Sentiment in Everyday News Page	+	−	+	○	○	−	+	−	−	−
3.3.	Sentiment in Daily News Production	+	−	+	−	○	−	+	−	−	−
3.4.	Sentiment in Vocabulary Use	+	−	+	−	+	−	+	−	−	−
Positive Inside, Negative Outside *?		Yes		Yes		No		Yes		No	

* Note: Asterisk means media builds positive image for home and negative image for world.

4. Discussion

Refer back to the two proposed research objectives: (1) What are the sentiment patterns in China's state-backed news websites today when reporting on domestic and international domains, as demonstrated on a day-to-day basis and supported by visualisation? (2) Can sentiment bias be confirmed using data from China's state-backed media itself, as well as through comparison with a benchmark media organisation? The results resolved the two questions using new collected data from four representative English news media outlets in China and the BBC as a benchmark. Sentiment patterns are extracted in terms of total news output, everyday sentiment level on news page, daily news production, and vocabulary using appropriate charts and graphs for visualisation. Even without the sentiment patterns from the BBC, evidence of bias from the four selected media websites in China speaks for itself. By comparing this with results from the BBC, the bias can be further highlighted and confirmed. The results demonstrate that sentiment bias exists in China's state-backed English-language news websites today, favouring positive coverage of the domestic side. Data collected on a day-to-day basis proved effective in bias identification, complementing the topical and issue-specific approach that has been used widely in previous research.

4.1. Three Sentiment Patterns and Natural Sentiment of News

Three sentiment patterns from the investigated news websites are identified in this study, according to the positivity and negativity of news titles under the domestic and world columns:

1. P1: Positive Inside, Negative Outside (CGTN, China Daily, Xinhuanet). This pattern reflects a typical bias and a typical strategy of China's state-backed media. Their sentiment in total news output, everyday news page, daily news production, and vocabulary usage around the home country is notably positive. When readers click

in the “China” column on the webpage on a day, they are more likely to see positive content than negative stories. From a news-generating and news-writing perspective, selecting positive news and using positive words when mentioning the home country is a deep-rooted habit. However, the patterns of “World” columns on these three websites are opposite. Although the “World” column of China Daily shows a more balanced daily production of positive and negative news, it still shifts towards negativity compared to its domestic counterpart.

2. P2: Neutral Inside, Negative Outside (Global Times). This pattern is trickier and more intriguing. Different from P1, it shows a more neutral or balanced news coverage about China. When readers navigate to the “China” column of Global Times and scroll, they have a roughly equal chance of seeing negative titles and positive titles. When titles are tagged with “China” or “Chinese”, vocabulary usage reflects the same typical tendency that P1 does. In terms of the “World” column, negative titles dominate. Overall, although the P2 pattern deviates from P1, it maintains more positive domestic coverage than positive foreign coverage.
3. P3: Negative Inside, Negative Outside (the BBC). As an external reference for studying China’s state-backed media, the BBC presents a sentiment pattern of negativity dominance. Negative titles prevail in terms of news webpages and production under both columns. In the network graph of the BBC, most big green (extremely positive) nodes can be found around its country and nationality nodes, but they only make up a small proportion; the number of negative nodes still surpasses that of positive ones. Visualisation allows this pattern to be grasped intuitively.

There are theories to explain the three types of sentiment patterns identified in this study. For P1, media control and authoritarian ruling philosophy can apply. As for P2, it can be considered a variation of P1, subtler and smarter. But the reason behind this difference is hard to know without knowledge of the news agency’s editing policy from insiders. In terms of the P3 negativity dominance pattern extracted from the BBC, the UK’s freer journalism, the BBC’s position in the organisational structure, and its editorial independence must be playing an important role.

Regarding how special the sentiment patterns of China’s state-backed media are, the BBC serves well as a benchmark or a point of comparison. To understand this further, it is worth discussing the big picture of news sentiment or the natural sentiment of news. Research has illustrated that negative news is prevalent in the United States ([Stone & Grusin, 1984](#); [Gieber, 1955](#); [Hartung & Stone, 1980](#)). [Shoemaker and Cohen \(2012\)](#) claimed that the nature of news is overwhelmingly negative. Such a perception of the negativity of news abounds. Plus, the proverb “No news is good news” embodies the same essence. It is perceived that negative messages are commonly and easily circulated. BBC news seems to be closer to the nature of negativity, while news under domestic columns of China’s media is not. The negative nature is driven by both the supply side (choices of editors and journalists) and the demand side (readers’ preference for negativity) of news media, according to [Soroka et al. \(2019\)](#). The hidden assumption of this theory is freedom of media and journalism, but China is far removed from such an environment because of its control, and that might be the reason why the negative nature is not observed when China’s media chooses what to put under its domestic columns. Following such logic, the domestic columns of CGTN, China Daily, and Xinhuanet seem to be tweaked intentionally. Even Global Times’s neutral coverage seems to be a sign of being controlled, but in a subtler manner.

4.2. Limitations

While bias was identified by analysing a half-year-long dataset from four of China's English media websites and highlighted by a smaller BBC dataset in an auxiliary way, limitations should be clarified. The first one is the imbalanced timeframe of the BBC dataset. If all five selected media sites had been tracked from the same start date and had yielded equivalent datasets, the conclusion would have been more robust. This limitation was partly addressed by the discussion on the negativity dominance of news in the Western context, which is well-supported by related research. Hence, the sentiment patterns of the BBC are almost certain to persist rather than reverse if the data collection timespan is extended.

The second point is about data collection timing and duration. Special events, crises, and national differences could cause sentimental fluctuations in news outlets. To observe overall sentiment patterns, it is worth asking, what timespan is long enough to mitigate the influence of those factors? Or is a half-year period long enough to represent media in general? For topic-centric empirical studies, corpora can be retrospectively gathered for many years or decades. Conversely, discovering patterns with data down to the day-to-day level usually requires tracking webpages in real time to obtain new data. Half a year is considered a good duration to represent a column in a news website for daily collection, as it allows for a sufficient number of special and normal events to emerge, escalate, and decline, representing the overall tendency. Although the longer the better, longer timespans have inherent challenges. Website scraping longer than half a year should raise a warning to researchers, because webpage layout, source code structure, and web accessibility may change, which requires more code modifications and leads to data discontinuity. An AI-powered web scraper (Khder, 2021) with high precision in news content recognition might be a better approach to make long-term monitoring easier in future research.

The third point is the use of a single news media outlet as the benchmark to highlight bias in the investigated Chinese English-language media. Admittedly, for holistic comparative news websites analysis, including multiple media from multiple countries is common (Starkey & Ye, 2017; Humprecht & Esser, 2018). The role of the BBC in this study was to highlight bias, and it served this purpose well since the result provided a straightforward and obvious contrast. Moreover, sentiment patterns in the four Chinese media outlets are strong and can stand in isolation. Nevertheless, a collection of English media sources from different countries with editorial independence, domestic/foreign columns, and a close governmental background to those media from China will enhance support for the conclusion.

Lastly, this study recognised bias but did not explore it in depth. The modest timeframe, domestic–foreign categorisation, and comparison with the benchmark are adequate for revealing bias. However, as mentioned before, a convincing topical analysis of news requires more retrospective data from a larger timespan and often relies on more diverse sources of news reports. A fine time granularity naturally conflicts with sound topical analysis. The combination of the two requires a more comprehensive study and more resources.

4.3. Future Work

This study showed the power of granular data frequency (day) for news bias (in sentiment) identification with a modest timeframe (half-year) and dichotomous categorisation (domestic–foreign). It should be acknowledged that discovering more nuanced bias requires data with finer-grained labelling (e.g., topics, issues) over a longer timespan (years, decades), with the analysis extended beyond sentiment (e.g., framing, narrative, readability), as mentioned in the limitations. Digital journalism is more dynamic than traditional journalism, and logging precise news timestamps is becoming increasingly

important. Responsiveness to events, delayed reporting, publication burst, story evolution, news persistence, and deletion are difficult to examine without fine-grained timestamp recording and constant monitoring. Future work will explore how to consolidate all the above considerations to build an integrated system with news collectors, databases, and dashboards. Such a system will allow researchers to investigate news in both time-series and topical dimensions granularly. Also, its advanced labelling and classification features are envisioned to enhance cross-media and cross-country comparisons to reveal more hidden trends and patterns.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/journalmedia6040177/s1>: Supplementary Materials S1: Project Dataset.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available in Supplementary Materials S1.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

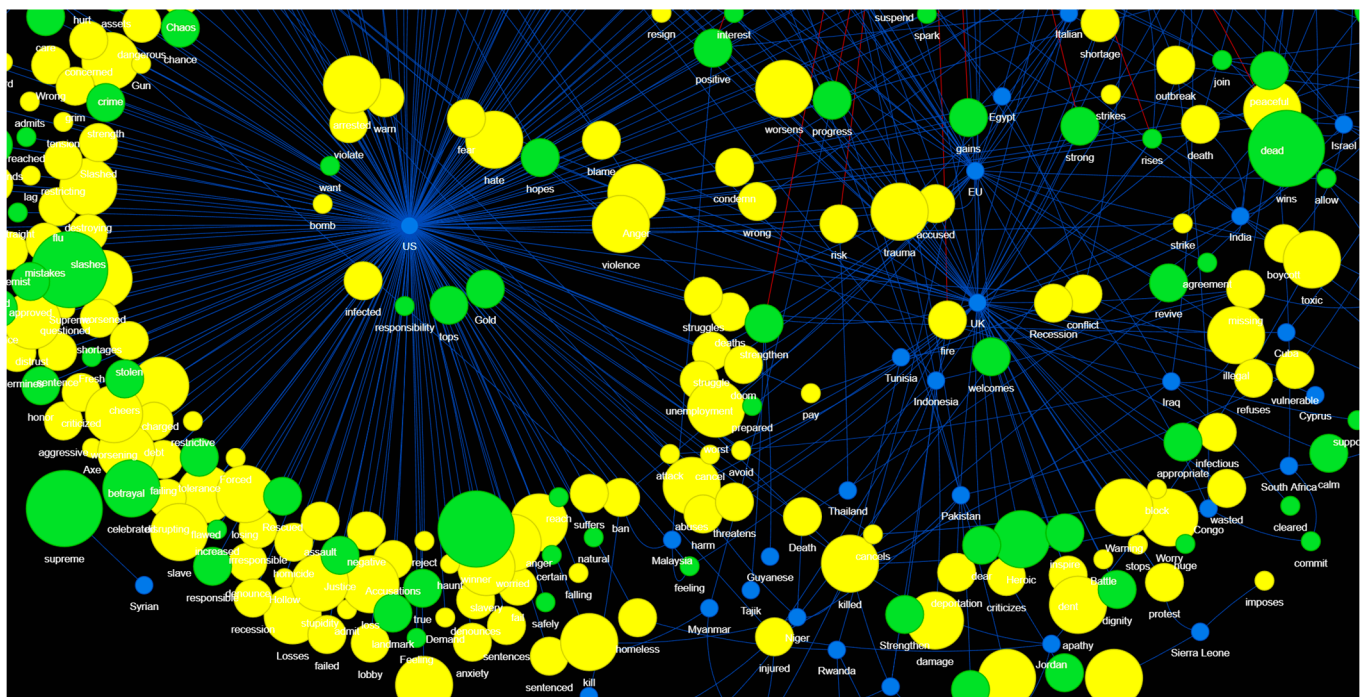


Figure A1. A closer look at the vocabulary network (China Daily). Negative token cluster. (Same visual encoding as Figure 14).

- Lagun, D., & Lalmas, M. (2016, February 22–25). *Understanding user attention and engagement in online news reading*. The Ninth ACM International Conference on Web Search and Data Mining (pp. 113–122), San Francisco, CA, USA.
- Liang, F. (2019). The new silk road on facebook: How China's official media cover and frame a national initiative for global audiences. *Communication and the Public*, 4(4), 261–275. [\[CrossRef\]](#)
- Lin, W. Y., Lo, V. H., & Wang, T.-L. (2011). Bias in television foreign news in China, Hong Kong, and Taiwan. *Chinese Journal of Communication*, 4(3), 293–310. [\[CrossRef\]](#)
- Liu, S., Guo, L., Mays, K., Betke, M., & Wijaya, D. T. (2019, November 3–4). *Detecting frames in news headlines and its application to analyzing news framing trends surrounding us gun violence*. The 23rd Conference on Computational Natural Language Learning (CoNLL) (pp. 504–514), Hong Kong, China.
- Montejo, G. M., & Adriano, T. Q. (2018). A critical discourse analysis of headlines in online news portals. *Journal of Advances in Humanities and Social Sciences*, 4(2), 70–83. [\[CrossRef\]](#)
- Morstatter, F., Wu, L., Yavanoglu, U., Corman, S. R., & Liu, H. (2018). Identifying framing bias in online news. *ACM Transactions on Social Computing*, 1(2), 1–18. [\[CrossRef\]](#)
- Nip, J. Y., & Sun, C. (2022). Public diplomacy, propaganda, or what? China's communication practices in the south China sea dispute on twitter. *Journal of Public Diplomacy*, 2(1), 43–68.
- Piotroski, J. D., Wong, T., & Zhang, T. (2017). Political bias in corporate news: The role of conglomeration reform in China. *The Journal of Law and Economics*, 60(1), 173–207. [\[CrossRef\]](#)
- Qin, B., Strömberg, D., & Wu, Y. (2014, June 6–10). *Media bias in autocracies: Evidence from China*. International Economic Association Conference, Dead Sea, Jordan.
- Qin, B., Strömberg, D., & Wu, Y. (2018). Media bias in China. *American Economic Review*, 108(9), 2442–2476. [\[CrossRef\]](#)
- Rieis, J., de Souza, F., de Melo, P. V., Prates, R., Kwak, H., & An, J. (2015, May 26–29). *Breaking the news: First impressions matter on online news*. The International AAAI Conference on Web and Social Media (Vol. 9, pp. 357–366), Oxford, UK.
- Samalis, A., Spyropoulos, A. Z., Makris, G. C., Bratsas, C., Veglis, A., Tsiantos, V., Baliou, A., Garoufallou, E., & Ventouris, A. (2023). Data journalism and network theory: A study of political communication through x (formerly twitter) interactions. *Journalism and Media*, 4(4), 1141–1168. [\[CrossRef\]](#)
- Schliebs, M., Bailey, H., Bright, J., & Howard, P. (2021). *China's inauthentic UK twitter diplomacy: A coordinated network amplifying PRC diplomats*. Oxford Internet Institute.
- Schweizer, D., Wang, X., Wu, G., & Zhang, A. (2025). Political connections and media bias: Evidence from China. *Journal of Corporate Finance*, 94, 102835. [\[CrossRef\]](#)
- Sheafer, T., & Dvir-Gvirsman, S. (2010). The spoiler effect: Framing attitudes and expectations toward peace. *Journal of Peace Research*, 47(2), 205–215. [\[CrossRef\]](#)
- Shen, F., & Guo, Z. S. (2013). The last refuge of media persuasion: News use, national pride and political trust in China. *Asian Journal of Communication*, 23(2), 135–151. [\[CrossRef\]](#)
- Shoemaker, P. J., & Cohen, A. A. (2012). *News around the world: Content, practitioners, and the public*. Routledge.
- Singh, G. (2016). Mass media in Xi's China: Markets versus control. *Strategic Analysis*, 40(5), 379–385. [\[CrossRef\]](#)
- Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, 116(38), 18888–18892. [\[CrossRef\]](#)
- Soroka, S., & McAdams, S. (2015). News, politics, and negativity. *Political Communication*, 32(1), 1–22. [\[CrossRef\]](#)
- Starkey, G., & Ye, H. (2017). Multimedia journalism: A comparative study of six news web sites in China and the UK. *GSTF Journal on Media and Communications*, 3(2), 1–10.
- Stockmann, D., & Gallagher, M. E. (2011). Remote control: How the media sustain authoritarian rule in China. *Comparative Political Studies*, 44(4), 436–467. [\[CrossRef\]](#)
- Stone, G. C., & Grusin, E. (1984). Network tv as the bad news bearer. *Journalism Quarterly*, 61(3), 517–592. [\[CrossRef\]](#)
- Tan, J., & Zhen, N. (2009). Chinese and Japanese newspaper reporting of the yasukuni shrine controversy: A comparative analysis of institutional media bias. *Electronic Journal of Contemporary Japanese Studies*, 9(1), 2.
- Tankard, J. W., Jr. (2001). The empirical approach to the study of media framing. In *Framing public life* (pp. 111–121). Routledge.
- Truex, R. (2016). Bias and trust in authoritarian media. *SSRN*. [\[CrossRef\]](#)
- Warren, J. (1988). Foreign and domestic news content of Chinese television. *Journal of Broadcasting & Electronic Media*, 32(2), 219–224. [\[CrossRef\]](#)
- Witzenberger, B., & Pfeffer, J. (2024). More inclusive and wider sources: A comparative analysis of data and political journalists on twitter (now x) in Germany. *Journalism and Media*, 5(1), 412–431. [\[CrossRef\]](#)

Xu, B., & Albert, E. (2014). Media censorship in China. *Council on Foreign Relations*, 25(1), 243–249.

Zhang, H., Zhou, S., & Shen, B. (2014). Public trust: A comprehensive investigation on perceived media credibility in China. *Asian Journal of Communication*, 24(2), 158–172. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.