

Article

Investigating COVID-19 News before and after the Soft Lock-down: An Example from Taiwan

Hsin-Yu Kuo ¹, Su-Yen Chen ^{2,*} and Yu-Ting Lai ³¹ Department of Educational Psychology and Counseling, National Tsing Hua University, Hsinchu 30013, Taiwan; kuohy@mx.nthu.edu.tw² Institute of Learning Sciences and Technologies, National Tsing Hua University, Hsinchu 30013, Taiwan³ International Intercollegiate Ph.D. Program (IPHD), National Tsing Hua University, Hsinchu 30013, Taiwan; yuting.estudio@gapp.nthu.edu.tw

* Correspondence: suychen@mx.nthu.edu.tw

Abstract: COVID-19 caused an unprecedented public health crisis and was declared a global pandemic on 11 March 2020, by the World Health Organization. The Taiwanese government's early deployment mitigated the effect of the pandemic, yet the breakout in May 2021 brought a new challenge. This study focuses on examining Taiwanese newspaper articles regarding the government response before and after the soft lockdown, collecting 125,570 articles reported by three major news channels from 31 December 2019, to 30 June 2021, and splitting them into four stages. Latent Dirichlet Allocation topic modeling and sentiment analysis were used to depict the overall picture of Taiwan's pandemic. While the news media focused on the impact and shock of the pandemic in the initial stage, prevention measures were more present in the last stage. Then, to focus on the government response indicators, we retrieved 31,089 related news from 125,570 news articles and categorized them into ten indicators, finding the news centered on the fundamental measures that were taken early and that were transformed into advanced measures in the latest and hardest period of the pandemic. Furthermore, this paper examines the temporal distribution of the news related to each indicator with the support of a sentiment analysis of the news' titles and content, indicating the preparation of Taiwanese society to confront the pandemic.

Keywords: COVID-19 pandemic; newspaper; government response; text mining; LDA; sentiment analysis; Taiwan

Citation: Kuo, H.-Y.; Chen, S.-Y.; Lai, Y.-T. Investigating COVID-19 News before and after the Soft Lock-down: An Example from Taiwan. *Sustainability* **2021**, *13*, 11474. <https://doi.org/10.3390/su132011474>

Academic Editor: John Rennie Short

Received: 5 September 2021

Accepted: 14 October 2021

Published: 17 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. 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 (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

COVID-19 has caused unprecedented global public health crises and was declared a pandemic on 11 March 2020, by the World Health Organization (WHO). The rapid spread of the virus, in turn, resulted in worldwide social and economic implications, with more than 182 million cases and 3.9 million deaths (<https://www.worldometers.info/coronavirus/>, accessed on 30 June 2021). Taiwan confronted the health crisis with surging infection rates, a total of 1210 cases and 12 deaths as of 11 May 2021, which further increased to 14,804 cases and 648 deaths as of 30 June 2021. While some countries had achieved high vaccination rates by and were lifting COVID-19 restrictions already in May 2021, Taiwan was still battling its worst outbreak with a so-called soft lockdown being imposed on 11 May 2021, as the government response was to shut down schools, bars and most public facilities, as well as allowing restaurants to only serve takeout.

Since the pandemic could have resulted in the breakdown of different dimensions of society [1], it is crucial to figure out how to maintain social functioning and minimize the damage caused by the crises from a sustainable development perspective. Analyzing the rich extant literature—including investigating the topics covered and their sentiments—can help us present a clearer picture of the pandemic and our social response to it. By

comparing the pre- and post-outbreak of the pandemic in May 2021, we can observe the fluctuation of positive and negative sentiments reflecting the public sentiment and then reflect on the crisis response that the Taiwanese government enacted. Compared to other relevant countries and jurisdictions, Taiwan had relatively more time to respond to the outbreak in 2021 and pinpoint the weaker aspects in the social systems.

Facing the formidable challenges of COVID-19, governments around the world have implemented a variety of policies in response to the pandemic. The European Union (EU) leaders agreed on the priorities of public health, travel and transportation, research and innovations, economy, crisis management, solidarity and education on 17 March and 20 March [2]. In the United States (US), the state government took the lead. California was the first state to impose a stay-at-home order as a response to COVID-19 [3]. To understand the actual versus appropriate timing of different government responses to mitigate their transmission and to investigate the effectiveness of government actions in reducing the number of deaths, large datasets have been compiled [4,5]. Moreover, few studies have reported on the variation of government responses [4–7], while some research papers have investigated the effectiveness of different governments [3,8–10]. Others have examined people's perceptions towards these responses, including some employing surveys as a research instrument [10–13] and some utilizing sentiment analysis and text mining [14–16].

In today's information society, people rely heavily on the Internet to better understand the global health crisis, with news media and social media being two of the channels. Most COVID-19-related sentiment analysis and text mining studies have analyzed social media data, while few comparative studies have suggested that the topic coverage of social media is narrower, i.e., the sentiment is more likely to be negative and has a shorter life span than that of news media [17,18]. By contrast, news articles that are written by journalists and subject matter experts present facts that make them more objective [18,19].

In this context, then, to better understand and gain more meaningful insight into the Taiwanese government response to the COVID-19 pandemic, this research paper explores the trends of its response indicators, as presented and reflected on news websites, using a news mining approach based on topic modeling and sentiment analysis. Then, we retrieve articles containing the terms related to government responses to conduct a sentiment analysis for a closer look, identifying positive, negative and neutral emotions as frames of news reporting and further inspecting the strategies employed.

2. Literature Review

2.1. Government Response

While modern societies have never confronted disease outbreaks of such scale and scope, the COVID-19 pandemic is today forcing governments around the world to make consequential policy decisions with limited information. To better understand the actual versus appropriate timing of government responses to mitigate their transmission—and to further investigate the effectiveness of different government actions in reducing the number of deaths—large datasets have been compiled and analyzed since the start of the pandemic. For example, the COVID-19 Government Response Event Dataset (CoronaNet v.1.0) has documented an extraordinary range of policies, including the type of policies, national and subnational enforcements, the group and geographical region targeted and the time frame implemented [4]. Employing topic modeling on records of over 8000 COVID-19-related government announcements from 190 countries between 31 December 2019, and 23 April 2020, with data from CoronaNet, researchers have found 13 themes, ranked in descending order of prevalence: external travel, health facilities, quarantine, tracking and testing, advisory systems, public awareness, nonessential businesses, government service, mass gatherings, school closure, curfew, health screening, state of emer-

gency and internal travel restriction. Research studies have also suggested that the variation in government responses was shaped by the opportunity learned from previous experience and the capacity to operationalize [6].

Similarly, the Oxford COVID-19 Government Response Tracker (OxCGRT) has documented policies and interventions from more than 180 countries since 1 January 2020, and provided a systematic set of cross-national and longitudinal measures, including 19 policy indicators [5]. This project tracked government responses across a standardized series of indicators and created a set of composite indices to measure the extent of the responses, covering the following: (1) closure and containment policies (C), such as school closing, workplace closing, cancellation of public events, restrictions on gathering size, public transport closing, stay-at-home, restrictions on internal movement and restrictions on international travel; (2) health policies (H), such as public information campaign, testing policies, contact tracing, emergency investment, investment in vaccines, facial coverage and vaccination policy; (3) economic policies (E), such as income support, debt relief, fiscal measures and giving international support; (4) others. More specifically, the OxCGRT research team examined the variations in government responses from 1 January to 1 September and found that, while some governments reacted immediately as the outbreak spread, others showed response lags [7]. Later on, with data analyzed from 1 January to 31 December, 2020, they also found that only a few countries had adopted strong closure and containment policies, i.e., lockdown policies, in early March; yet, within a month, intensive responses had become a global phenomenon, with many countries lifting and re-imposing policies in a response see-saw as the epidemic went up and down. Furthermore, on the one hand, among the 19 indicators, there was a 40% chance that countries had carried out public information campaigns, international travel restrictions and testing policies within 10 days after the first requirement policy, showing closure or containment and health policies established before economic support. On the other hand, as policies eased up, even though closure or containment policies were loosened, health and economic policies were maintained [5]. Regarding the effectiveness of these government response indicators, utilizing global data from OxCGRT between March and September 2020, a study reported that quick and early action by the government in imposing strict measures was important in slowing down the spread of the virus and lowering fatalities, particularly regarding closure and public awareness policies [8].

Finally, researchers from different academic fields also examined whether specific government interventions should be employed or intensified. For example, applying reduced-form econometric methods with data on 1700 interventions imposed across China, South Korea, Italy, Iran, France and the US, a global-level study found that closure or containment policies—such as business or school closures, travel bans and bans on social gatherings, as well as other types of social distancing—had prevented over 61 million cases across these six countries by the end of March 2020 [8]. Moreover, another study used a model based on the Verhulst equation to assess the impact of the total versus partial lockdowns in China, Italy and Spain from January to April 2020. The results suggested that the infected population size at lockdown time, showing the rate of new infection not being out of control, seemed to play a significant role in the infection spread, regardless of the lockdown type [10]. A third study investigated the relationships among three policies, i.e., shutdowns, reopening and mask-wearing, and the COVID-19 infection growth rates in the US from February to August 2020 and recommended a short period of shutdown followed by a reopening, albeit with the universal mask covering being continued [3].

2.2. People's Perceptions towards Government Response

While aforementioned government responses concerning containment and mitigation policies had been implemented in many European states around mid-March 2020, there were continued debates and discussions in every society regarding the appropriateness of the measures taken. To better understand these people's perceptions, a survey of

over 7000 representative samples in seven countries, i.e., Denmark, France, Germany, Italy, Portugal, the Netherlands and the United Kingdom (UK), was conducted in the first two weeks of April 2020. The issues included school closures, bans on public gatherings, border closures, bans imposed on the export of medical equipment, fines for quarantine violations, random temperature checks, curfews, public transport suspensions and tracking cases and their contacts with mobile phone data. At a more general level, the results suggested that government responses toward COVID-19 were well received by the population in all countries. The most approved measures were fines for quarantine violations, bans for public gatherings and border closures, each supported by 83% of the population. On the other hand, measures related to individual freedom and privacy, such as tracking and curfews, were disapproved by 23% of the population. In addition, people seemed to be worried about the balance between saving lives and saving livelihoods, especially those from southern Europe, which signaled that effective strategies of communication were significant for securing a sufficient level of trust regarding policymakers and a high level of compliance [11]. Similarly, the willingness to be vaccinated against COVID-19 was surveyed in seven European countries, with a sample size of 7662 in April 2020. On average, 73.9% of the respondents reported that they would be willing to be vaccinated, ranging from 80% in the UK and Denmark, to 62% in France. The most pronounced reason for hesitating or not wanting to be vaccinated was concerning its side effects [12]. In Asia, in turn, a survey was also conducted to investigate the Japanese citizens' behavioral changes in response to the government's precautionary measures against COVID-19, with 11,342 participants, by the end of March 2020. As of March 28, 1499 cases and 49 deaths had been confirmed in Japan, whereas a total of 571,678 cases and 26,494 deaths had been reported globally. These findings indicate that 86.6% of the respondents reported practicing social distancing by avoiding mass gatherings, while 70.1% reported always wearing a surgical-style facemask when going out. On the other hand, gender, age, income, marital status and personality were found to be significant predictors for not conducting social distancing. To be more precise, being male, younger, not married, from a low household annual income and being out-going were more likely not to follow preventative government policies related to social distancing [13].

In addition to surveys as an instrument of research, sentiment analysis has been another popular methodology employed by academic researchers to explore people's perceptions towards their government COVID-19 responses. For instance, a Spanish study has identified affective tones on confinement measures, opinions related to social distancing, self-isolation and quarantine, from the main media and digital ecosystem, Twitter, YouTube, Instagram, official press websites and Internet forums, with a text corpus of 80,091. The data was collected from 9 March to 1 May, 2020, and was divided into three distinctive stages. These findings revealed four main topic groups, namely, COVID-19 illness numbers, concerns over lockdown, the uncertainty of the situation and its consequences and others (such as gratitude and donations), as well as a pattern of polarization on the topic of isolation, regarding the highest peak of anger around the beginning of stage 2, and a steady increase in joy during stage 1, which stabilized during stage 2 and 3. Furthermore, there were some events showing the synergies of feelings, such as joy and sadness, joy and fear, and anger and fear, demonstrating social media as a form of "collective therapy" [14]. Second, social-media-user opinions on remote work during the pandemic—as a chain effect of workplace closing, concerning the economic consequences and worrying about the new normality—were also critically investigated, with English tweets between 1 February and 10 August, 2020, and tweets from March 2019 utilized for selective comparative analysis. Results indicated that the topic of remote work at the epidemic peak in March 2020 having increased almost 15 times; 62.23% of the posts were positive while 12.96% were negative [15]. Third, digital polarization of mask-wearing in the US during COVID-19 was investigated with an examination of Twitter posts from 1 March to 1 August, 2020. From a total of 35 distinct hashtags, 74% were found to be pro-mask, while, from the 412,959 tokens of hashtags, 93.6% were found to be pro-mask [16].

2.3. Sentiment Analysis and Text Mining with Social and News Media

Text mining and sentiment analysis studies have also employed news media as a data source, either with other types of data sources, or as a single source. For example, one study compared sentiment in relation to social media and government press and another study in relation to social and news media during the pandemic. First, a study monitored the digital ecosystems during March and April 2020 in Spain, with 106,261 communications from the perspective of public crisis communication. As for social media, a multi-layered perceptron was trained with a set of communications created by the Spanish Society for Natural Language Processing which contained over 100,000 natural language texts tagged with a positive versus negative message for each communication. As for the government press, a content analysis of all press releases during this period was also conducted. These findings suggest that the messages emitted by the Spanish government have always wanted to give a positive tone in their communications, with none referring to either the infected or the dead. On the contrary, the feelings of the public regarding COVID-19 have generated peaks in different emotions, revealing that they were mixed with negative sentiments—among which sadness, disgust, anger and fear—for reasons including party criticism of the government responses, contradictions of the medical experts, the country's status being one of the most affected counties in the world and confinement suffering [20]. In a similar vein, another study compared news and social media in Portuguese from January to May 2020 in Brazil by employing the sentiment analysis and topic modeling approaches. With 18,413 news articles from the main news media website in the country, i.e., Universo Online, researchers found that the top themes were politics, prevention and control, whereas, with 1,597,934 tweets posted by 1,299,084 users, the top theme was stories, focusing on personal opinions and cases. Moreover, news media indicated that all themes were more positioned around neutral polarity, yet those of social media were distributed lower on the sentiment scale, with politics, prevention, control and confirmed cases around the negative polarity [18].

Regarding text mining and sentiment analysis with news media alone, first, a study examined how news media in China delivered health communication during the early stages of the pandemic, with 7791 articles collected from one of the most reputable Chinese media content datasets, i.e., the WiseSearch, from 1 January to 20 February, 2020. By employing the method of Latent Dirichlet Allocation (LDA), researchers chose 20 as the number of topics and then categorized them into nine main themes, with the top three on prevention and control procedures, medical treatment and research, and global social and economic influences, accounting for 60% in total. Findings pointed out that the Chinese news media lagged behind the development of the pandemic by focusing on the whole society without fulfilling personal needs at the same time, i.e., instructions on individual prevention, clinic choices and detection [21]. Second, politicization and polarization regarding COVID-19 news in US newspapers and television networks from March to May 2020 were compared, with an initial database of 36,620 stories, yet only 6985 were analyzed in which the keyword of COVID-19 was mentioned at least three times for substantive coverage on the topic. Results showed that, while both media are highly polarized, newspaper coverage is more political than TV network news, for the latter not only covered politicians but also scientists. The researchers pointed out that media coverage appeared to contribute to the polarization of public attitudes towards COVID-19 [22]. Third, an international comparison was carried out to investigate COVID-19 news coverage across four nations, i.e., UK, India, Japan and South Korea, with more than 100,000 news headlines and articles collected from 1 January to 1 December, 2020, from the English language websites of eight major newspapers from these counties. The researchers employed the Top2Vec model on the datasets for each country. With 23,821 articles, the UK dataset produced 308 topics, with the top five being about maternity during COVID-19, education, Australia-related news, US-related news and the Office for National Statistics (ONS). With 47,342 articles and 402 topics, India's top five were Cricket league and IPL, US, Punjab, law and order, and vaccine development. As for Japan, with 21,039 articles generating

255 topics, the country's top five topics were global stock exchange, Nikkei, Tokyo 2020 Games postponement, COVID-19 cases and South Korea. As for South Korea, with 10,076 articles and 127 topics, the country's top five topics were economic relief, geopolitics, impact on retail, lower GDP and COVID-19 cases. To conclude, topic modeling revealed common themes to be education, economy, US and sports across four nations and the sentiment analysis indicated that the UK had 73.23% of news being negative, in contrast to South Korea's 54.47% being positive [19]. Fourth, based on media framework theory, the study of Thirumaran et al. [23] focused on newspapers from Singapore and New Zealand and investigated the relationship between destination crisis management strategy and the effects of news portrayal out of traveling concerns. It was further maintained in the research paper that the success or failure of crisis management in a certain area would make the sentiment embedded in the media of that area inclined towards being positive or negative and further influence local reputation.

According to the discussion above, sentiment analysis is usually applied to social media to reveal people's feelings and perspectives towards certain measures during the pandemic [14–16] and examining how social media influences risk communication [20], as well as comparing the polarity of different kinds of media [22]. When applying sentiment analysis to cross-national comparison of news articles, news-related sentiments can reflect the impact on the countries during the pandemic and, also, the effectiveness of their crisis management more generally [19,23].

2.4. Taiwan in Global Context

While many countries adopted similar measures to respond to the enormous impact caused by the COVID-19 outbreak, Taiwan conducted early deployment to mitigate the effects of the pandemic [24]. Taiwan Centers of Disease Control (CDC) has played a vital role in responding to the COVID-19 pandemic. Having learned from the experience of Severe Acute Respiratory Syndrome (SARS) in 2003, the Taiwanese government realized that it was indeed crucial to establish a well-organized framework to control infectious diseases. During the pandemic, deploying systems such as quarantine, digital fencing and name-based mask distribution could successfully reduce the risk of community infection [25]. Nevertheless, in Taiwan, these preventative measures had to have a law basis, to be under the supervision of citizens and to be followed with four principles, i.e., rapid measures, early deployment, prudent actions and transparency.

The government response in Taiwan included surveillance and laboratory diagnosis, border control, control of community transmission, medical system response and preparedness, stockpile and allocation of PPE and other medical supplies, health education, fighting disinformation and loosening epidemic prevention measures [26]. The temporal patterns of the spread of COVID-19 (both worldwide and in Taiwan) and the corresponding measures taken by the WHO and Taiwanese government, respectively, are shown in Figure 1. The number of new daily cases shown is the rolling seven-day average (<https://ourworldindata.org/covid-deaths>, accessed on 30 June 2021). The WHO adopted many response measures, yet we only extracted some of them during the initial stage of the pandemic. Border control policies and national testing network have been gradually implemented and established since January 2020. The mask distribution system was implemented in February 2020. The “Special Act for Prevention, Relief and Revitalization Measures for Severe Pneumonia with Novel Pathogens” plan was announced in February 2020, including supporting medical personnel, relief, subsidies, compensation and revitalization measures, and punishment for disseminating false information and violating the isolation measures.

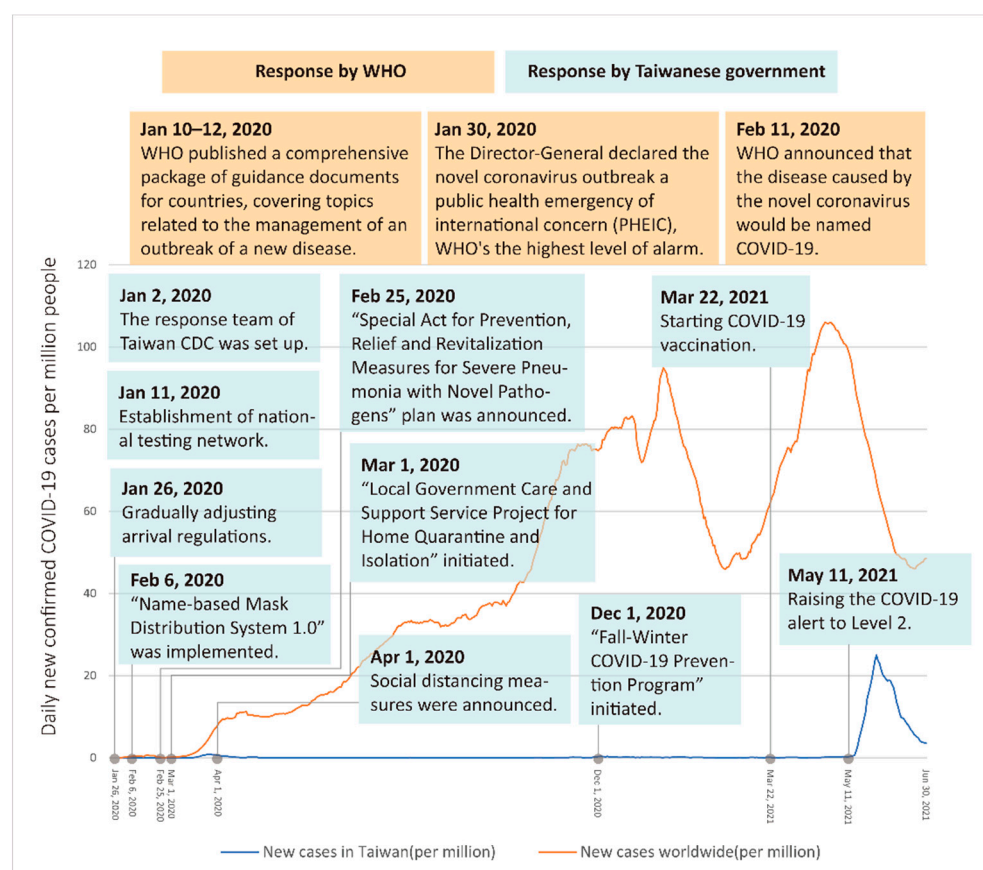


Figure 1. Spread of COVID-19 and response measures adopted by the Taiwanese government.

To sum up, according to the rich cluster of literature reviewed above, some of the studies examined several key measures implemented within a single country [3,13,14,16,20], whereas other studies compared certain issues internationally [10–12]. The methods included survey and text mining. Three common research materials of text mining seem to come from social media, government publications and newspapers, examined by means of headline and content analysis. These text mining approaches, in turn, can be applied to have a more comprehensive understanding of the content conveyed by the media [18–22]. More importantly, in the same context, text media framing can be observed and detected through sentiment analysis [19]. The proportion of positive and negative words can reveal the journalists' intention to convey important public messages and, more generally, the situation of whether the public attitude is positive or negative towards the pandemic.

Within the above framework, the current study takes news articles in Taiwan as its main research object. In so doing, it comprehensively explores news content by treating sentiments as a media frame based on sentiment analysis [23]. Then, it critically examines the overall and specific aspects of Taiwanese news relevant to government response indicators.

3. Research Methodology

3.1. Research Questions

Based on the literature reviewed above, this study hereby aims to analyze the topics and sentiments expressed in online news and compare them in different time periods, during the COVID-19 pandemic in Taiwan. As part of this, we formulated four relevant research questions:

- RQ1: What topics appear in the Taiwanese news articles and how have they changed before and after the soft lockdown?
- RQ2: What sentiments are expressed in online news articles before and after the soft lockdown?
- RQ3: How were government response indicators covered and distributed before and after the soft lockdown?
- RQ4: What sentiments expressed in online news articles correspond to government response indicators before and after the soft lockdown?

3.2. Data Sampling

We selected three major Taiwanese newspapers: United Daily News, Liberty Times and China Times. We scraped 180,249 articles with keywords such as COVID-19, pandemic, or coronavirus in Mandarin Chinese from December 2019 to June 2021. The collected data included hyperlinks, dates, news titles and news content. Due to the fact that there were many articles mentioning COVID-19, yet less relevant to the pandemic, we filtered the articles by the principle that COVID-19-related keywords were mentioned at least three times in the content to be included for analysis [22]. Finally, 125,570 articles were compiled as the dataset used for analysis. We employed CkipTagger, an open-source library built by CKIP [27], for word segmentation and syntactic parsing for preprocessing [28].

According to the events unfolding regarding the pandemic and related regulation announcements, we divided the data into four stages based on the development of COVID-19 in Taiwan:

- Stage 1, from 31 December 2019, to 31 May 2020. During the initial stage, more and more new confirmed cases were reported worldwide, yet the characteristics of the virus were still relatively unknown. In Taiwan, the first confirmed case was reported on 28 January. The first confirmed death was announced on 16 February. The primary prevention measures were announced and implemented from January to April. National and international case numbers continued to rise. Then, however, the number of locally confirmed cases decreased to zero in mid-April. The pandemic had eased at the end of Stage 1.
- Stage 2, from 1 June 2020, to 30 November 2020. During this stage, the economic activities resurged along with the easing of restrictions. Some measures have remained, such as wearing a mask in public spaces, e.g., train stations, and prohibiting international travel.
- Stage 3, from 1 December 2020, to 30 April 2021. During this stage, the so-called “Fall–Winter COVID-19 prevention program” was initiated on 1 December 2020, and included border quarantine, community prevention and medical response. A locally confirmed case was reported on December 22 and broke the record of no-new locally confirmed cases for 253 days straight. Four new locally confirmed cases were reported on 12 January 2021, causing temporary panic. At the end of stage 3, i.e., in April, a quarantined hotel had violated the basic quarantine rules, leading to the disease spreading uncontrollably from a hotel.
- Stage 4, from 1 May 2021, to 30 June 2021. After the hotel spread of Stage 3, the pandemic was out of control, with a low vaccine rate of less than 1% of the population as of 1 May 2021. The Central Epidemic Command Center (CECC)—an agency of the National Health Command Center—announced the raising of the COVID-19 alert level to 2 on May 11 and to 3 on May 19, thereby implementing a so-called soft lockdown and strengthening the preventative measures more generally. The number of confirmed cases and deaths increased in mid-May, yet they had again decreased by June of the same year.

3.3. Latent Dirichlet Allocation (LDA) and Model Evaluation

LDA is a particularly popular and representative algorithm for fitting a topic model and a generative probabilistic model of a corpus. Its basic idea is that documents are represented as random mixtures over latent topics, where a distribution over words characterizes each topic [29]. Newspapers usually include several topics at the same time. LDA treats each document as a mixture of topics and each topic as a mixture of words. This allows documents to “overlap” each other in terms of content, rather than being separated into discrete groups, in a way that mirrors the typical use of natural language.

The variable names are defined to describe the model and generative process [29–31] as follows: a document denoted by d is a sequence of N_d words which are the basic units of discrete data; a corpus denoted by D is a collection of M documents ($d \in \{1, \dots, M\}$); K is the number of topics, θ_d is a distribution of topics in document d and φ_k is a distribution of words in topic k . LDA assumes the following generative process for each document d with length N_d in a corpus D :

1. Choose a multinomial distribution φ_k for topic k ($k \in \{1, \dots, K\}$) from a Dirichlet distribution with parameter β .
2. Choose a multinomial distribution θ_d for document d ($d \in \{1, \dots, M\}$) from a Dirichlet distribution with parameter α .
3. For a word w_n ($n \in \{1, \dots, N_d\}$) in document d :
 - a. Choose a topic z_n from θ_d .
 - b. Choose a word w_n from φ_{z_n} .

Only the generated words in the documents are the observed variables in the above process. Referring to the latent variables φ and θ and hyperparameters α and β maximizes the joint probability of the words in a document of the corpus. Then, we can finally determine the latent topics for each document according to the estimated θ_d .

For RQ1, we applied LDA from the Gensim package, an open-source library for unsupervised topic modeling and natural language processing and selected the hyperparameter value as $\alpha = \text{“auto”}$, $\eta = \text{“auto”}$, $\text{passes} = 10$ and $\text{iterations} = 50$. To determine the number of topics, we used the numbers from 5 to 30. By obtaining the topics—10 words corresponded to the topics and two quantitative indicators, namely, the perplexity and coherence scores, through the LDA algorithm—we evaluated the number of topics based on the coherence score and the meaning of each topic. The performance of the model is better when the coherence score is higher [31]. Then, we selected 20 as the optimal number of topics. Finally, we assigned the topics for each document according to the highest probability of the topic occurring in the document.

3.4. Sentiment Analysis

For RQ2, we calculated sentiment scores for news titles and content by applying a lexical approach [32]. We followed two steps. First, we used CkipTagger to identify the part of speech for each word in a sentence based on syntactic parsing. Second, we computed the sentiment score using the previous results and augmented the NTU sentiment dictionary (ANTUSD) [33,34] for each sentence and document. The words were assigned a sentiment score according to the part of speech and ANTUSD. The summation of word sentiment scored from the score of a sentence; the score of a sentence converted to 1 if it was more than 1 and to -1 if it was less than -1. Afterwards, the document sentiment scores consisted of the summation of the sentence scores, which were then adjusted by the weights based on the length of the sentences. We adjusted the dictionary to fit the context of the COVID-19 pandemic to obtain the result more accurately. For instance, “increase” is generally positive, yet not in the pandemic context in newspapers, especially for reporting COVID-19 cases. The sentiment scores of documents were between -1 and 1 and we transformed the scores to “positive” if they were more than or equal to 0.5, “negative” if less than or equal to -0.5 and “neutral” if between 0.5 and -0.5.

3.5. Identification of the Government Response Indicators in News

In addition, to focus on the government response (RQ3), we retrieved 31,089 news from 125,570 articles of the dataset. The process is here described. According to the Oxford COVID-19 government response tracker (OxCGRT) [5] and the main measures in Taiwan [24–26], we identified 10 indicators that were common subjects in Taiwanese newspapers: testing policy (H2), contact tracing (H3), facial coverings (H6), vaccination policy (H7), school closing (C1), workplace closing (C2), cancellation of public events (C3), restrictions on gathering size (C4), restrictions on international travel (C8) and income support (E1). A total number of 31,089 news titles included these keywords. Then, we adopted the news titles which contained specific keywords assuming that, if the terms appeared in the news title, they were more likely to be keywords in the content. For some of the titles that had more than one keyword about government response indicators, we employed the statistic frequency-inverse document frequency (TF-IDF) to find the importance of certain words and differentiate the indicators of the titles. With the assistance of the technique, a Gensim package for the TF-IDF was employed. The TF-IDF value increases depending on the term frequency in the document and is offset by the number of documents that contain the term in the corpus [35]. We took the ones with the highest TF-IDF value as the classification keywords for the 10 government response indicators.

Moreover, for RQ4, we calculated the sentiment scores of titles and contents and distinguished three sentiments, positive, negative and neutral, as the above principle. Although it is unnecessary to talk about the government response when these words appear in the news titles, it could be argued that at least a certain extent of sentiment would be generated in readers' minds that further influence the public's attitudes and feelings towards such measures.

4. Results

4.1. Topic Distribution across Four Stages

For classification in news articles, LDA is useful as a fast-filtering algorithm for feature selection, providing an efficient way to find hidden topics in a large-scale data collection without manual tagging. In our dataset with 125,570 articles, we applied LDA in the news content. Examining the different numbers of topics, we identified 20 as optimal and extracted the latent topics. We obtained the top five keywords for each topic by the algorithm and named the topics according to the meaning of keywords. Assigning the topic to each article according to probability computed by LDA, we could observe the prevalence of the topics. The topic names, article numbers and the top five keywords corresponding to each topic are shown in Table 1.

Table 1. Topic classification and keywords.

Topic No.	Topic	N(%)	Top 5 Keywords
1	Impact on the economy	25,824 (21%)	Chinese, pandemic, impact, market, economy
2	Individual prevention in Taiwan	18,127 (14%)	People, prevention, pandemic, measure, government
3	Confirmed cases worldwide	14,827 (12%)	Confirmed, pandemic, pneumonia, patient, die
4	Institutional prevention in Taiwan	13,771 (11%)	Staff, work, central, government, Institution
5	Information on the pandemic	12,596 (10%)	Pandemic, Facebook, COVID-19, Beijing, hope
6	Confirmed cases in Taiwan	11,291 (9%)	Symptom, fever, case, report, contact

7	Global trend	9235 (7%)	US, Japan, country, hygiene, government
8	Research on coronavirus	6665 (5%)	Infect, coronavirus, COVID-19, discovery, expert
9	Economic relief packages in Taiwan	2656 (2%)	Relief, industry, package, apply, finance
10	Vaccine policy	2045 (2%)	Development, go abroad, our side, immunity, clinical
11	Medical personnel in Taiwan	1935 (2%)	Staff, treatment, medical, testing, patient
12	Border control	1697 (1%)	Quarantine, isolate, enter a country, airport, home
13	Event cancellation in Taiwan	1170 (1%)	Event, cancel, aviation, tourism, hold
14	Information technologies	803 (1%)	Service, provide, technology, through, system
15	School closing	755 (1%)	School, student, take care, bank, ministry of education
16	Impact on economic livelihood	648 (1%)	Operation, sale, legal person, customer, business
17	Local disinfection	618 (0%)	Staff, alcohol, clean, go to work, company
18	Combating misinformation in Taiwan	458 (0%)	Message, man, hotel, police, buy
19	Online learning and services	385 (0%)	Online, video, child, ministry of labor, learn
20	Pandemic in Taipei markets	64 (0%)	Ko Wen-Je, Taipei city, Taipei, wholesale, market

The first six topics were impact on the economy, individual prevention in Taiwan, confirmed cases worldwide, institutional prevention in Taiwan, information on the pandemic and confirmed cases in Taiwan. By assigning the topic to each article according to probability computed by LDA, we could observe the prevalence of the topics. More than 20% of the articles discussed the impact on the economy, while 14% concentrated on individual prevention in Taiwan and 11% on institutional prevention, which meant that around a quarter of articles talked about prevention in Taiwan. Then, 12% of the articles reported on confirmed cases worldwide, while 9% reported on confirmed cases in Taiwan. In total, 10% were about the information of the pandemic, usually mentioning information channels, such as Facebook. The topics about economic relief packages in Taiwan, vaccine policy and border control related to government responses accounted for 2%, 2% and 1% of the dataset, respectively.

To reveal the temporal relationship between the prevalence of topics in the news articles and the number of confirmed cases, we analyzed them and the results are presented in Figure 2. We divided the periods into four distinct stages based on the status of the epidemic in Taiwan. At the beginning of the outbreak, due to the uncertainty of the pandemic trend and the rapid increase in the number of confirmed cases, the whole world was in shock and the first confirmed case of COVID-19 was announced in Taiwan. Therefore, what followed was that a great number of news articles were published in a short time period. According to the results of topic modeling, Taiwanese news articles mainly focused on the impact on the economy, confirmed cases worldwide and global trends, as well as highlighted individual and institutional prevention in Stage 1. After May 2020, in Stage 2, since the number of domestic confirmed cases had been effectively under control,

the focus of the news turned to the impact on the economy and confirmed cases in other countries. At the end of November 2020, the Taiwan Centers for Disease Control (CDC) announced the Fall–Winter COVID-19 Prevention Program, which was going to be launched in December, including mask-wearing policy in the eight types of high-risk venues, reinforcement of case reporting and specimen collecting in medical institutions and border inspection. There were no domestic cases until 20 December 2020, 253 days since the last domestic case. On 12 January 2021, two domestic confirmed cases were reported. In Stage 3, the articles discussed institutional prevention early, yet they were centered on the impact of the economy and confirmed cases worldwide after February 2021. While border control was loosened in March 2021 (in Stage 3), the case of an airline pilot was confirmed and the pandemic started to be out of control again. On 11 May 2021, the CDC announced that Taiwan had entered the stage of community infection and raised the pandemic warning level to 2. In the meanwhile, the number of news articles regarding individual and institutional prevention and confirmed cases in Taiwan was significantly increasing in Stage 4.

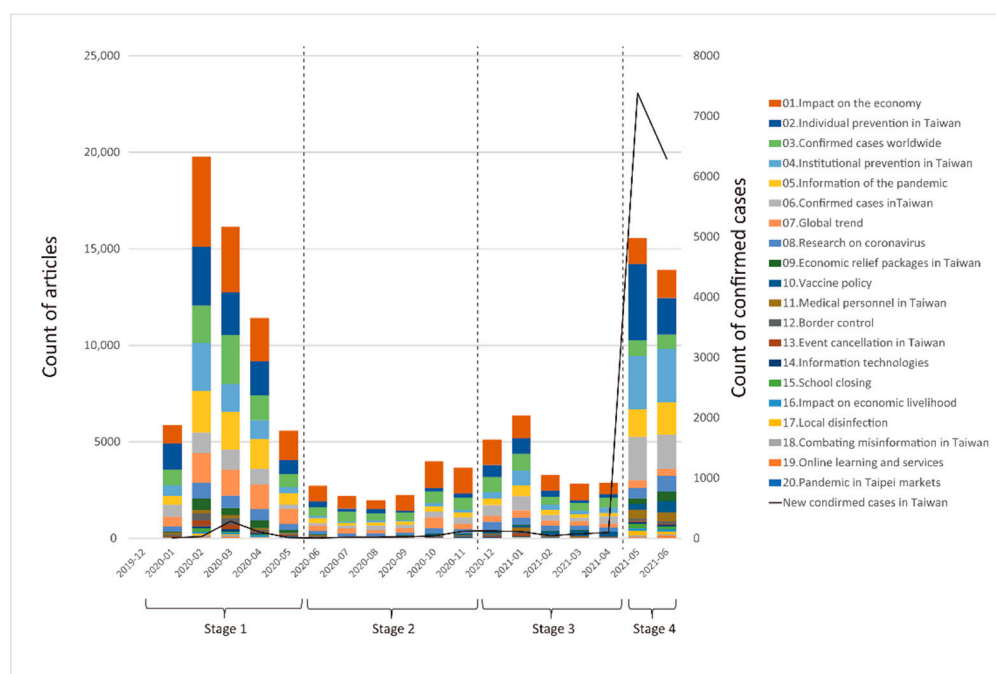


Figure 2. Prevalence and temporal variation of topics.

4.2. Sentiment Distribution across Four Stages

By assigning each news title and content a sentiment score and converting it to positive, neutral and negative, we can observe the sentiment expressed by the news articles as shown in Figures 3 and 4. The number of neutral sentiment articles is the most common, both in titles and content. The sentiment trend was negative in news titles but positive in news content. This study utilizes the lexicon-based approach and dictionary-based methods. The sentiment scores depend on the number of positive and negative words in the document.

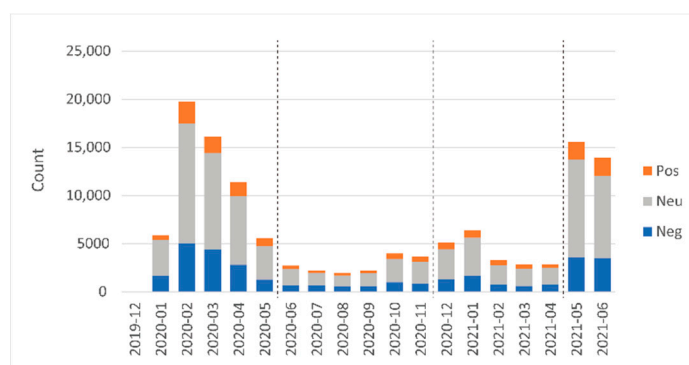


Figure 3. Sentiment distribution in news titles.

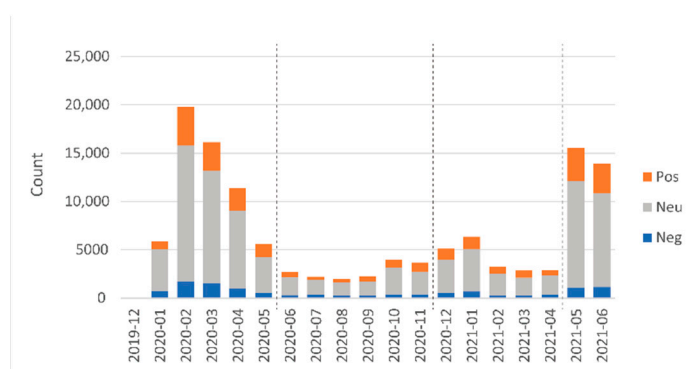


Figure 4. Sentiment distribution in news contents.

The proportion of negative sentiments was higher than that of positive sentiments in the news titles. The high peak of negative sentiment was in February 2020. By contrast, the proportion of positive sentiments was higher than that of negative sentiments in news content. The negative titles and content were both more common in Stage 1. The second increase in the number of negative news titles took place in Stage 4 after the soft lockdown.

The news titles were designed to catch the attention of the public and used more negative terms. Thus, the sentiment scores tended to be negative, especially reporting confirmed deaths in Stages 1 and 4. In the early days of Stage 3, there were two major events which resulted in COVID-19 infections, which, in turn, caused a slight increase in negative news titles. The news content often mentions terms such as “execute” and “implement”, the aim being for the public to notice the government’s prevention rules and thereby the sentiment scores of the contents also tended to be positive.

4.3. Government Response Indicators Distribution across Four Stages

We provide an overall picture in Sections 4.1 and 4.2 of this paper for our dataset with 125,570 news articles, which include COVID-19-related keywords that appear at least three times in the content. Although topic modeling can identify different topics, those topics can contain several measures. For example, individual prevention in Taiwan may include issues such as face masks, crowd gathering and avoiding going to hospitals. Therefore, it is difficult to echo a specific government response directly. Then, to retrieve news articles regarding the primary government responses, we selected keywords of them and used the TF-IDF to compute the weight of these keywords in the news title. In total, 31,089 news titles included the keywords of government response indicators. As shown in Figure 5, the dataset’s composition of government response indicators accounts for 15%–22% in Stage 1, 15%–23% in Stage 2, 21%–39% in Stage 3 and 36%–44% in Stage 4, which implies that the newspapers reported more information on response measures. We identified the government response indicators in 31,089 news titles. For individual

indicators, the temporal distribution is shown in Figure 6. “Facial coverings” and “restrictions on international travel” were frequently discussed in Stage 1. The number of articles on “vaccination policy” increased during Stages 2 and 3. In Stage 4, all articles regarding government response indicators, except “restrictions on international travel”, increased after the soft lockdown and reports centered on “vaccination policy” and “testing policy”. “Income support”, in turn, increased during Stages 1 and 4.

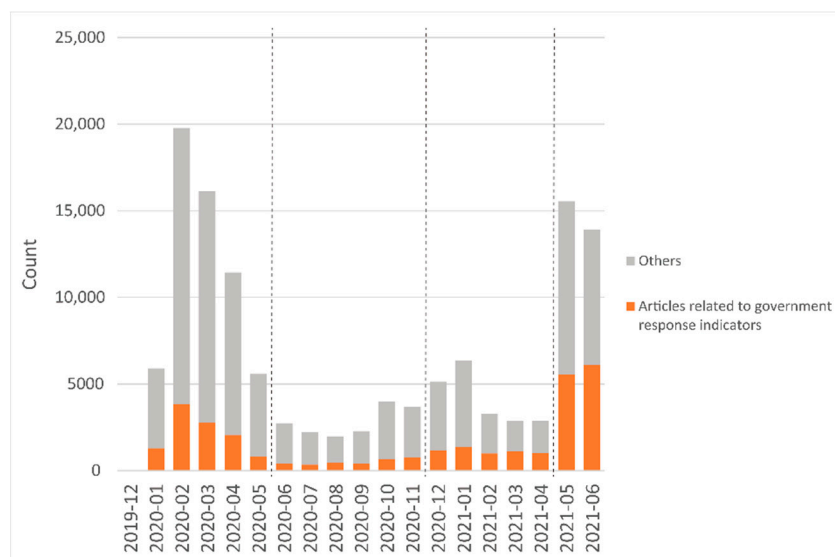


Figure 5. Prevalence and temporal variation of articles related to the government response indicators in dataset.

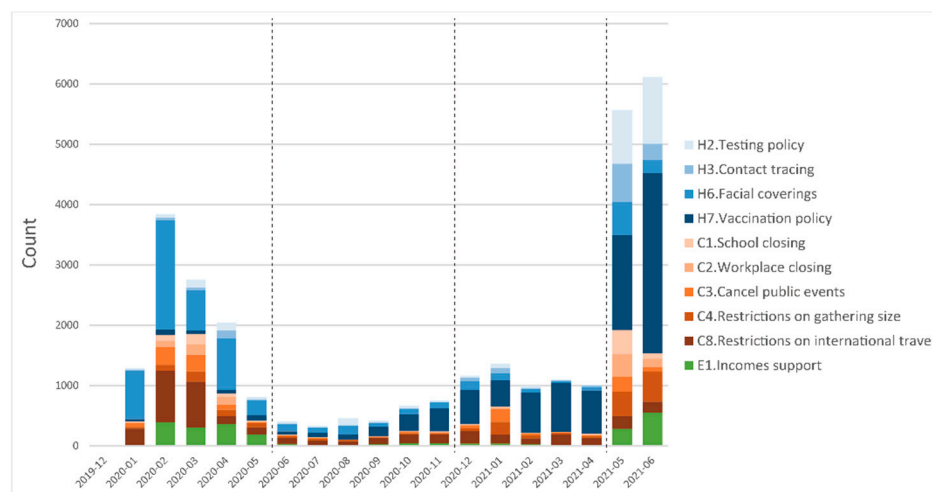


Figure 6. Prevalence and temporal variation of articles related to government response indicators.

4.4. Government Response Indicator Sentiments across Four Stages

The government response indicators in news titles correspond to the sentiment and the temporal distribution is shown in different patterns in Figures 7–10. Figure 7 shows the pattern of gaining high attention in Stage 1 and then a general decreasing tendency in the following stages, exemplified by “facial covering” and “restrictions on international travel”. The proportions of negative and positive sentiments of “facial covering” were similar in news titles, yet, in news content, the proportion of positive sentiments was higher. The news titles with negative sentiments in “restrictions on international travel” were more common than positive ones, yet they were similar in news contents.

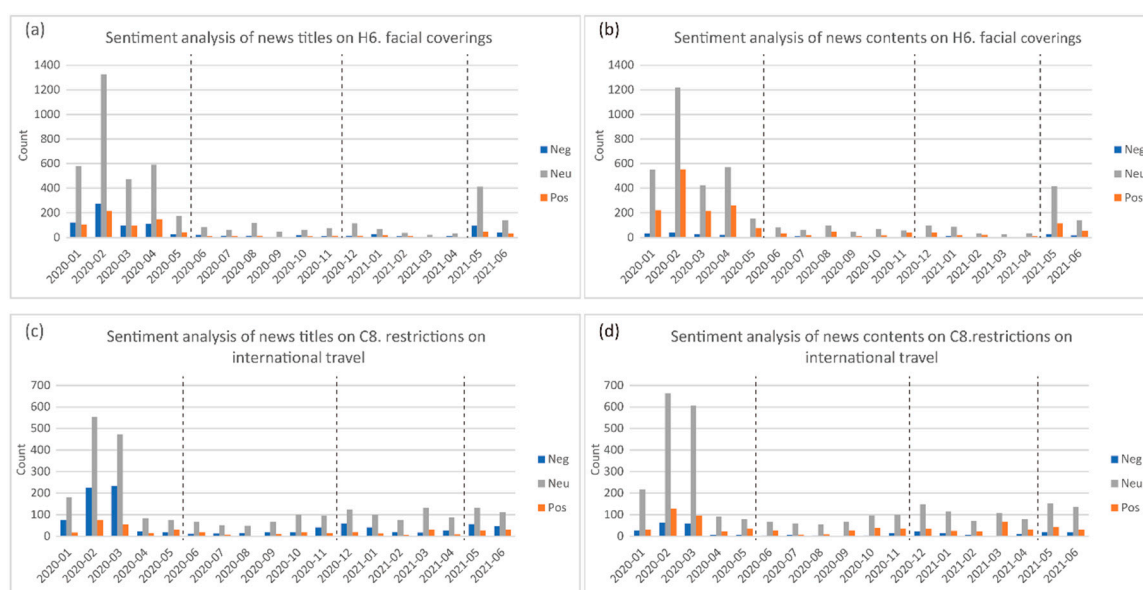


Figure 7. Patterns of increase in Stage 1: sentiment analysis of (a) news titles on H6. facial coverings, (b) news contents on H6. facial coverings, (c) of news titles on C8. restrictions on international travel, and (d) news contents on C8. restrictions on international travel.

By contrast, “contact tracing”, “testing policy” and “vaccination policy” were less mentioned in Stage 1, yet they increased in the other stages, as shown in Figure 8. “Contact tracing” and “testing policy” both increased in Stage 4. The increase in “contact tracing” depended on the infection rates more generally. The CECC and local governments announced the tracking of confirmed patients and people would then notice whether or not their tracks overlapped at the same time. The CECC enforced a real-name registration system in Stage 4. Different from the overall trend, the positive news titles regarding “testing policy” were more common than the negative ones, for more reports centered on someone passing a test and COVID-19 rapid testing kits for domestic use could be easily bought. The topic of “vaccination policy” started increasing from Stage 2 onward and surged in Stage 4. For the lack of vaccines, low immunization coverage and worry regarding the side effects of vaccines, there were more negative news titles than positive ones in Stage 4.

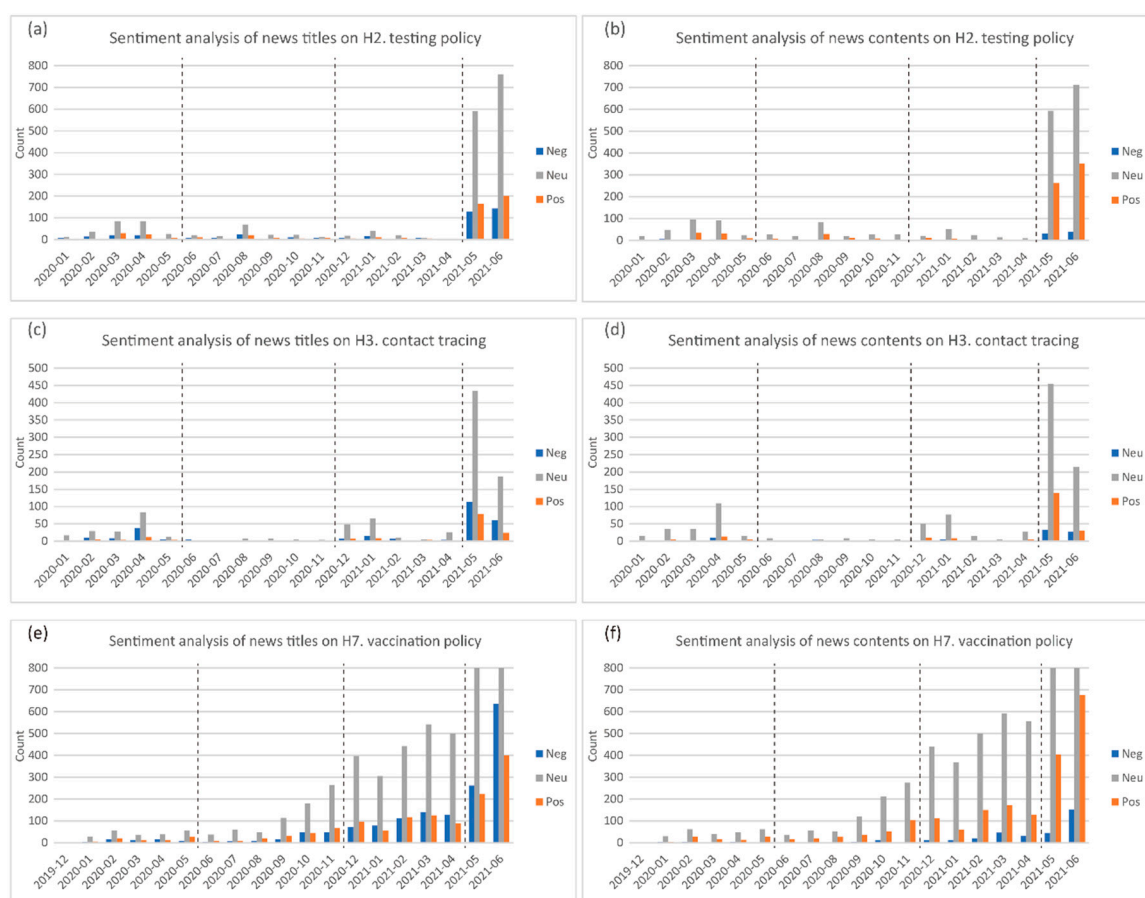


Figure 8. Patterns of increase after Stage 1: sentiment analysis of (a) news titles on H2. testing policy, (b) news contents on H2. testing policy, (c) news titles on H3. contact tracing, (d) news contents on H3. contact tracing, (e) news titles on H7. vaccination policy, and (f) news contents on H7. vaccination policy.

The four government response indicators, i.e., “school closing”, “workplace closing”, “cancellation of public events” and “restricts on gathering size”, increased during Stages 1 and 4, as shown in Figure 9. The patterns of “school closing” and “workplace closing” were similar in that they achieved a high peak after the first month of soft lockdown. However, they were slightly different in terms of sentiments in Stage 4. The proportion of positive articles was higher than that of negative ones in terms of “school closing”, but the opposite in reference to “workplace closing”. Many resources regarding online learning and the skills for online classes were reported in newspapers with positive sentiments. In addition to the popularization of hardware and software in most areas in Taiwan, the reaction to “school closing” and “workplace closing” dropped in later days of Stage 4. “Cancellation of public events”, in turn, increased along with the reports appearing regarding the announcement of event cancellation or postponement in early periods of infection outbreaks. When the virus began to spread, “restrictions on gathering size” also increased and cases caused by gathering were reported along with negative sentiments.



Figure 9. Patterns of increase in Stage 1 and Stage 4: sentiment analysis of (a) news titles on C1. school closing, (b) news contents on C1. school closing, (c) news titles on C2. workplace closing, (d) news contents on C2. workplace closing, (e) news titles on C3. cancel public events, (f) news contents on C3. cancel public events, (g) news titles on C4. restrictions on gathering size, and (h) news contents on C4. restrictions on gathering size.

The pattern of “income support”, as shown in Figure 10, is similar to Figure 9, in that the reports centered on the government response indicator in Stages 1 and 4. Yet, the proportion of positive news seemed higher both in news titles and contents, for the news usually used terms with positive sentiments, such as relief, subsidy and support.

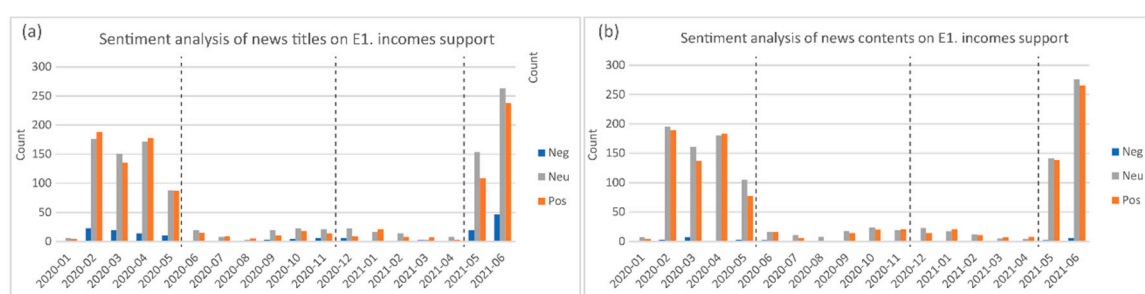


Figure 10. Patterns of increase in Stage 1 and Stage 4 along with the positive sentiments: sentiment analysis of (a) news titles on E1. incomes support, and (b) news contents on E1. incomes support.

5. Discussion

Newspapers inform citizens of important societal events, including the impact of coronavirus, government responses and the results of these responses. Topic modeling and sentiment analysis have been applied widely in previous academic studies to better understand what and how the news media convey information to the public [18–23]. Within this similar framework, this study collected 125,570 news articles from three major news channels in Taiwan from 31 December 2019, to 30 June 2021, and divided this period into four stages for RQ1 and RQ2. Then, we selected 31,089 news articles whose titles included keywords related to government response indicators for RQ3 and RQ4. The key findings regarding each research question are discussed below.

5.1. RQ1: What Topics Appear in the Taiwanese News Articles and How Have They Changed before and after the Soft Lockdown?

We applied LDA topic modeling to find latent topics in the news contents. We evaluated the model and chose 20 as the number of topics, recognizing the topic of each document based on the probability of each topic occurring in it. As shown in Table 1 and Figure 2, the distribution of article numbers and topic proportion was different in the four stages. In the initial period of the pandemic in Stage 1, the news media focused on topics regarding its impact on the economy, personal and institutional prevention in Taiwan and confirmed cases worldwide. In Stages 2 and 3, the world was already being profoundly influenced by the pandemic. Taiwan's news media followed the same trend and reported on the impact of the economy and confirmed cases worldwide continuously. The prevention measures were further strengthened in winter in early Stage 3, with COVID-19 still globally prevailing. After the soft lockdown in Stage 4, due to the worsening pandemic in Taiwan, the news media enhanced the attention of the locals, especially regarding prevention. In this Stage, differently from Stage 1, the government had built some systems with sufficient materials and regulations to respond to pandemic prevention requirements. The highest number of monthly news articles related to COVID-19 was February 2020, the next was March 2020 and then May 2021, with the highest confirmed case rates. To conclude, the newspapers centered on the impact of the pandemic in the previous stages and prevention measures in the later stages with relatively detailed information on COVID-19.

5.2. RQ2: What Sentiments Are Expressed in Online News Articles before and after the Soft Lockdown?

Adopting the lexicon-based approach and dictionary-based methods, we show the sentiment polarity expressed in the news titles and contents; the distributions are shown in Figures 3 and 4. Unlike social media [15–18], sentiments are more likely to be neutral in the news to present the facts. Negative news titles and positive contents increased after the soft lockdown. The proportions of positive titles and negative content were low. A probable reason for this is that idiomatic words and sentence structures are different in news titles and content. The first and second highest months of negative sentiment, both in terms of titles and contents, were February and March 2020, which did not respond to the highest number of confirmed cases. The result responds to the variation in topics in the news which emphasize the impact or prevention.

5.3. RQ3: How Were Government Response Indicators Covered and Distributed before and after the Soft Lockdown?

A high coverage rate of government response indicators was in Stage 4 with the highest numbers of confirmed cases. The highest number of monthly news articles related to

COVID-19 was February 2020. Figure 5 shows the temporal distribution of indicators related to the articles. In Stage 1, “facial coverings” and “restrictions on international travel” were popular. “Vaccination policy” occurred in Stage 2 and remained in Stages 3 and 4. Apart from “restrictions on international travel”, most articles on indicators increased in the early period of Stage 4. The count on news articles was in partial line with the requirement of people and government responses in different stages. In addition to the results of LDA topic modeling, we provided a second interpretation and focused on government response indicators as keywords. While, in the previous stages, the reports emphasized fundamental prevention measures, such as face coverings and border control, we experienced more advanced measures, such as testing, contact tracing and vaccines in the later period of this Stage.

5.4. RQ4: What Sentiments Expressed in Online News Articles Correspond to Government Response Indicators before and after the Soft Lockdown?

Based on the above results and sentiment analysis on news titles and contents, we organized the distribution of articles on 10 indicators and inducted them into four different patterns. Figure 7 depicts the process of shaping mature fundamental measures, including “facial coverings” and “restrictions on international travel”. Figure 8 shows the implementation and improvement of advanced measures indicators containing “contact tracing”, “testing policy” and “vaccination policy” in Stage 4. The increase in “contact tracing” depended on the infection events, for the CECC and the local governments announced the tracks of confirmed patients to people, so they could find out whether or not their tracks overlapped. The CECC enforced a real-name registration system and related rules in Stage 4. Different from the overall trend, the positive news titles regarding “testing policy” were more common than negative ones due to the fact that the majority of reports were focused on people passing the test and the availability of the rapid domestic COVID-19 testing kits on the market. “Vaccination policy” started to increase from Stage 2 onward and surged in Stage 4. The phenomenon of more negative news titles appearing in Stage 4 was the result of several reasons, such as lack of vaccines, low immunization coverage and concerns regarding the side effects of the vaccines, further implying insufficient communication between the government and the public on vaccination policy.

Indicators in Figure 9 and Figure 10 include “school closing”, “workplace closing”, “cancellation of public events”, “restrictions on gathering size” and “income support”, which have more relationship to daily life and livelihood, rather than to medical-related prevention indicators. The patterns of “school closing” and “workplace closing” were similar in that they both achieved a high peak after the first month of soft lockdown. However, they were slightly different in terms of sentiments in Stage 4. The proportion of positive sentiments is higher than the negative one for “school closing”, since the news provided information on online learning resources with positive words. By contrast, the phenomenon is reversed in “workplace closing”. Besides students taking classes online and many workplaces changing the mode of work or sales, the functions of schools and workplaces did not stop. “Cancellation of public events” increased along with the appearance of news on the announcement of events cancellation or postponement during the early periods of the pandemic outbreak. When the outbreak started, “restrictions on gathering size” were rising and cases caused by gathering were reported along with the negative sentiment. The pattern of “income support” shown in Figure 10 is similar to Figure 9—the reports centered on the government response indicator in Stages 1 and 4. Nevertheless, even during the hardest times, the proportion of positive news still seemed higher in both news titles and content, given that the news usually used terms with positive sentiments, such as relief, subsidy and support. Examining the indicators by the pattern with sentiment variation, we found that the discrimination of sentiment of news titles was better than content. Only the negative sentiments of content on “restrictions on gathering size” was more common than the positive ones. “Testing policy”, “school closing” and “income

support” imply that positive sentiments are more common than negative ones, indicating that the measures were actively promoted by the government and citizens.

Compared to studies that use text mining on news data in other areas [19,21], most news media seemed to be inclined to report more on the impact of COVID-19 on societies, but less on prevention at an individual level. Although Taiwan’s news media focused on the economy early, they also provided information regarding primary individual prevention based on government responses. They tracked the global pandemic when the epidemic was slowing down in Taiwan, indicating that the Taiwanese government not only referenced the knowledge against SARS early, but it also borrowed the experiences from other countries and jurisdictions from Stages 1–3. Unlike social media that reflect the feelings of the public with strong sentiments [14–16], newspapers present knowledge objectively, including government responses, the development and influence of pandemic and the result of preventative measures. The highest monthly number of overall news was in Stage 1, yet the highest number and proportion of the news related to indicators was in Stage 4. Synthesizing the results of the temporal distribution of subjects classified by LDA and government response indicators revealed that the news media and society could focus more on policy indicators in the latest stages of the pandemic.

6. Conclusions

The COVID-19 pandemic could be considered as one of the world’s most important and far-reaching recent events. Taiwan’s principles of pandemic prevention were mostly practical, especially when information regarding the virus was scarce. However, the breakout in May 2021 brought a new challenge. The newspapers recorded the pandemic from various angles and its impact on society. To maintain social functioning, we should survey the weaknesses and advantages of government response to crises from a sustainable viewpoint. As part of this, newspapers provide valuable information to citizens of countries and regions. The findings of this study, in turn, provide an overall picture of the Taiwanese news coverage and government response indicators, as detailed above. This paper reviewed government responses to the spread of the virus and people’s perceptions towards these responses. It looked at the sentiment analysis and text mining of social and news media during the COVID-19 pandemic. In particular, we have critically examined the pandemic by placing Taiwan within a global context, indicating that deploying systems based on a well-organized framework is crucial to prevent and control the spread of the disease.

We examined 125,570 news articles from 31 December 2019, to 30 June 2021, in Taiwan, finding 20 topics in the news by LDA topic modeling methods and polarities of news titles and contents by sentiment analysis. The prevalence of topics regarding prevention and locally confirmed cases and the number of negative news titles became higher when the pandemic worsened. Furthermore, we chose 10 government responses as indicators and observed them in the newspapers filtering the news articles by the topic including the keywords of government responses. The fluctuation in news count related to the response indicators shows the characteristic of news coverage in the early period of soft lockdown and with a different focus in four stages. Two tendencies can be observed. First, a focus on the impact of the pandemic has shifted into a focus on prevention measures when the outbreak started in Taiwan. Second, news articles related to government response indicators progressed from fundamental measures, including “facial covering” and “restrictions on international travel” to advanced measures, including “contact tracing”, “testing policy” and “vaccination policy”. Then, we examined the patterns of news titles and content related to the indicators and considered the sentiments, indicating the “facial coverings” and “restrictions on international travel” as fundamental measures that received less attention during Stage 4, while “contact tracing”, “testing policy” and “vaccination policy” received a lot of public attention after lockdown, with “vaccination policy” receiving considerable negative sentiment.

In comparison with the findings from other studies, we conclude this study as follows. First, regarding the improvement of government responses, the research study by Capano et al. showed that factors such as the opportunity, capacity of each government, the nature of national leadership, the organization of government and civil society and blind spots towards the vulnerabilities helped shape policy responses across nations to the pandemic [6]. We supply a perspective that borrowing from the experiences of other countries was beneficial for shaping government responses in the global context. Second, from the viewpoint of people's perceptions towards government responses, an investigation indicated that one of the critical issues was balancing saving lives and saving livelihoods, in the EU, and the pandemic may have acted as a stressor causing health and economic anxieties [11]. The impact on the economy was also a concern in Taiwan and the response of income support was quick to reduce the impact on economic livelihood. Third, by comparing previous research papers on news media conducted in other countries, a study employed news headlines for sentiment classification across four countries, namely, the UK, India, Japan and South Korea, and they only focused on general topics, such as education, economy, US and sports [19]. Rather, we employed both news titles and contents and focused on government response indicators in the pandemic, providing social implications to conquer the challenge of the pandemic. To sum up, in addition to people's perceptions, news data analysis helped us examine our social functioning and society's preparation when facing a disaster such as the COVID-19 pandemic and improve the responses towards a sustainable society.

The main contributions of this study include the following. First, we presented how Taiwan prepared for and made progress on fighting the pandemic through analyzing data retrieved from news articles. In the initial stages, a shock regarding the start of the epidemic accounts for a high proportion in the newspapers. As we further inspected the content of the government response indicators, it indicated that some basic prevention measures, such as mask-wearing and border control, were gradually accepted, acknowledged and even supported by the public. It is worth noting that we still saw news tracking the global pandemic trend and referring to the experience of other countries and jurisdictions, even at the time when the pandemic was slowing down in Taiwan, i.e., during Stages 2 and 3. In Stage 4, with the highest number of confirmed cases and deaths, our research shows that the influence and shock caused by the pandemic did not come under the spotlight. Instead, the news put more emphasis on preventative measures. Moreover, the proportion of government response indicators significantly increased in Stage 4. After the soft lockdown was introduced during this stage, the focus transferred to more advanced preventative measures, such as testing, contact tracing and other containment policies. The results illustrate that Taiwan was well prepared to a certain degree and was capable of maintaining sustainable societal development during the turbulent pandemic era.

Second, we contributed some applications with the text mining approach. LDA and sentiment analysis of news titles and contents on comprehensive data helped us obtain a more complete picture of the news more generally. By focusing on the coverage of government response indicators, we identified articles by title, which included keywords of government responses indicators, with the help of TF-IDF. Furthermore, analyzing temporal sentiments for each indicator enabled us to gain beneficial insights into the overall government responses to the spread of the virus and people's perceptions towards these responses.

Third, we investigated the results of government response indicators by observing the temporal distribution of the sentiments of news titles and content, which was performed separately. Four patterns were then observed, showing the different degrees of discussion in four stages, indicating the degree of preparation for each indicator. Most indicators were accepted by the public, albeit the vaccine policy was and remains controversial in Taiwan.

There are three main limitations to this study. First, we analyzed the news in Taiwan within a global context and it could not make inferences to other societies. Furthermore, we selected only 10 out of 19 Oxford COVID-19 government response tracker indicators based on the experience in Taiwan. Some indicators, such as stay-at-home requirements, could not be identified in the Taiwanese news. Third, we used a lexicon-based approach for sentiment analysis to identify the sentiments for each text unit automatically. Although we adjusted the dictionary to fit the pandemic context, certain sentiments of documents could not be accurately expressed because of polysemy.

As for suggestions for future academic research in this study field, we could argue that the threat of the pandemic has not ceased globally; therefore, conducting cross-cultural research on newspapers is perhaps more valuable now than it has ever been in the past. Within the same framework, collecting and synthesizing data with a hybrid approach such as ours, i.e., questionnaire surveys and social media analysis, could provide further beneficial insights into our understanding of the overall government responses to pandemics and people's perceptions towards these responses.

Author Contributions: Conceptualization, S.-Y.C.; data curation, H.-Y.K. and Y.-T.L.; formal analysis, H.-Y.K., S.-Y.C. and Y.-T.L.; funding acquisition, S.-Y.C.; methodology, H.-Y.K. and S.-Y.C.; writing—original draft, H.-Y.K., S.-Y.C. and Y.-T.L.; writing—review and editing, S.-Y.C. and Y.-T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research study was funded by the Ministry of Science and Technology in Taiwan under the grant number MOST 110-2511-H-007-006-MY2.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be provided on request by the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. OECD. Building a Coherent Response for a Sustainable Post-COVID-19 Recovery. *OECD Policy Responses to Coronavirus (COVID-19)* **2020**, 1–28, doi:10.1787/d67eab68-en.
2. Goniewicz, K.; Khorram-Manesh, A.; Hertelendy, A.J.; Goniewicz, M.; Naylor, K.; Burkle, F.M., Jr. Current Response and Management Decisions of the European Union to the COVID-19 Outbreak: A Review. *Sustainability* **2020**, *12*, 3838, doi:10.3390/su12093838.
3. Zhang, X.; Warner, M. E. COVID-19 Policy Differences across US States: Shutdowns, Reopening and Mask Mandates. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9520, doi:10.3390/ijerph17249520.
4. Cheng, C.; Barceló, J.; Hartnett, A. S.; Kubinec, R.; Messerschmidt, L. COVID-19 Government Response Event Dataset (CoronaNet v.1.0). *Nat. Hum. Behav.* **2020**, *4*, 756–768.
5. Hale, T.; Angrist, N.; Goldszmidt, R.; Kira, B.; Petherick, A.; Phillips, T.; Webster, S.; Cameron-Blake, E.; Hallas, L.; Majumdar, S.; Tatlow, H. A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **2021**, *5*, 529–538.
6. Capano, G.; Howlett, M.; Jarvis, D. S. L.; Ramesh, M.; Goyal, N. Mobilizing Policy (in)Capacity to Fight COVID-19: Understanding Variations in State Responses. *Policy Soc.* **2020**, *39*, 285–308.
7. Hale, T.; Petherick, A.; Phillips, T.; Webster, S. Variation in government responses to COVID-19' version 4.0. Blavatnik School of Government Working Paper. Available online: <http://www.bsg.ox.ac.uk/covidtracker> (accessed on 7 April 2020).
8. Chisadza, C.; Clance, M.; Gupta, R. Government Effectiveness and the COVID-19 Pandemic. *Sustainability* **2021**, *13*, 3042.
9. Hsiang, S.; Allen, D.; Annan-Phan, S.; Bell, K.; Bolliger, I.; Chong, T.; Druckenmiller, H.; Huang, L.Y.; Hultgren, A.; Krasovich, E.; et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nat. Cell Biol.* **2020**, *584*, 262–267.
10. Sanchez-Caballero, S.; Selles, M. A.; Peydro, M. A.; Perez-Bernabeu, E. An Efficient COVID-19 Prediction Model Validated with the Cases of China, Italy and Spain: Total or Partial Lockdowns? *J. Clin. Med.* **2020**, *9*, 1547, doi:10.3390/jcm9051547.
11. Sabat, I.; Neuman-Böhme, S.; Varghese, N. E.; Barros, P. P.; Brouwer, W.; van Exel, J.; Schreyögg, J.; Stargardt, T. United but Divided: Policy Responses and People's Perceptions in the EU during the COVID-19 Outbreak. *Health Policy* **2020**, *124*, 909–918.
12. Neumann-Böhme, S.; Varghese, N. E.; Sabat, I.; Barros, P. P.; Brouwer, W.; van Exel, J.; Schreyögg, J.; Stargardt, T. Once We Have It, Will We Use It? A European Survey on Willingness to Be Vaccinated against COVID-19. *Eur. J. Health Econ.* **2020**, *21*, 977–982, doi:10.1007/s10198-020-01208-6.

13. Muto, K.; Yamamoto, I.; Nagasu, M.; Tanaka, M.; Wada, K. Japanese Citizens' Behavioral Changes and Preparedness against COVID-19: An Online Survey during the Early Phase of the Pandemic. *PLOS ONE* **2020**, *15*, e0234292, doi:10.1371/journal.pone.0234292.
14. Iglesias-Sánchez, P. P.; Vaccaro Witt, G. F.; Cabrera, F. E.; Jambrino-Maldonado, C. The Contagion of Sentiments during the COVID-19 Pandemic Crisis: The Case of Isolation in Spain. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5918, doi:10.3390/ijerph17165918.
15. Wrycza, S.; Maślankowski, J. Social Media Users' Opinions on Remote Work during the COVID-19 Pandemic. Thematic and Sentiment Analysis. *Inf. Syst. Manag.* **2020**, *37*, 288–297, doi:10.1080/10580530.2020.1820631.
16. Lang, J.; Erickson, W. W.; Jing-Schmidt, Z. #MaskOn! #MaskOff! Digital Polarization of Mask-Wearing in the United States during COVID-19. *PLOS ONE* **2021**, *16*, e0250817, doi:10.1371/journal.pone.0250817.
17. Kim, E. H.-J.; Jeong, Y. K.; Kim, Y.; Kang, K. Y.; Song, M. Topic-Based Content and Sentiment Analysis of Ebola Virus on Twitter and in the News. *J. Inf. Sci.* **2016**, *42*, 763–781, doi: 10.1177/0165551515608733.
18. de Melo, T.; Figueiredo, C. M. S. Comparing News Articles and Tweets about COVID-19 in Brazil: Sentiment Analysis and Topic Modeling Approach. *JMIR Public Health Surveill.* **2021**, *7*, e24585, doi: 10.2196/24585.
19. Ghasiya, P.; Okamura, K. Investigating COVID-19 News across Four Nations: A Topic Modeling and Sentiment Analysis Approach. *IEEE Access* **2021**, *9*, 36645–36656, doi: 10.1109/ACCESS.2021.3062875.
20. de Las Heras-Pedrosa, C.; Sánchez-Núñez, P.; Peláez, J. I. Sentiment Analysis and Emotion Understanding during the COVID-19 Pandemic in Spain and Its Impact on Digital Ecosystems. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5542, doi:10.3390/ijerph17155542.
21. Liu, Q.; Zheng, Z.; Zheng, J.; Chen, Q.; Liu, G.; Chen, S.; Chu, B.; Zhu, H.; Akinwunmi, B.; Huang, J.; Zhang, C. J. P.; Ming, W. K. Health Communication through News Media during the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach. *J. Med. Internet Res.* **2020**, *22*, e19118, doi:10.2196/19118.
22. Hart, P. S.; Chinn, S.; Soroka, S. Politicization and Polarization in COVID-19 News Coverage. *Sci. Commun.* **2020**, *42*, 679–697, doi:10.1177/1075547020950735.
23. Thirumaran, K.; Mohammadi, Z.; Pourabedin, Z.; Azzali, S.; Sim, K. COVID-19 in Singapore and New Zealand: Newspaper Portrayal, Crisis Management. *Tour. Manag. Perspect.* **2021**, *38*, 100812.
24. Tan, T. W.; Tan, H. L.; Chang, M. N.; Lin, W. S.; Chang, C. M. Effectiveness of Epidemic Preventive Policies and Hospital Strategies in Combating COVID-19 Outbreak in Taiwan. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3456, doi:10.3390/ijerph18073456.
25. Chen, S.-C. Taiwan's Experience in Fighting COVID-19. *Nat. Immunol.* **2021**, *22*, 393–394.
26. Taiwan Centers for Disease Control. Prevention and Control of COVID-19 in Taiwan. Available online: https://www.cdc.gov.tw/en/category/page/0vq8rsAob_9HCi5GQ5jH1Q (accessed on 26 August 2021).
27. CkipTagger. Available online: <https://github.com/ckiplab/ckiptagger> (accessed on 26 August 2021).
28. Chinese Knowledge and Information Processing, Academia SINICA. CKIP Transformers. Available online: <https://ckip-transformers.readthedocs.io/en/stable/main/readme.html> (accessed on 26 August 2021).
29. Blei, D. M.; Ng, A. Y.; Jordan, M. I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
30. Jelodar, H.; Wang, Y.; Yuan, C.; Feng, X.; Jiang, X.; Li, Y.; Zhao, L. Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey. *Multimed. Tools Appl.* **2019**, *78*, 15169–15211.
31. Paek, S.; Kim, N. Analysis of Worldwide Research Trends on the Impact of Artificial Intelligence in Education. *Sustainability* **2021**, *13*, 7941, doi:10.3390/su13147941.
32. Alamoodi, A. H.; Zaidan, B. B.; Zaidan, A. A.; Albahri, O. S.; Mohammed, K. I.; Malik, R. Q.; Almahdi, E. M.; Chyad, M. A.; Tareq, Z.; Albahri, A. S.; Hameed, H.; Alaa, M. Sentiment Analysis and Its Applications in Fighting COVID-19 and Infectious Diseases: A Systematic Review. *Expert Syst. Appl.* **2021**, *167*, 114155, doi:10.1016/j.eswa.2020.114155.
33. Ku, Lun-Wei, Ho, Xiu-Wei and Chen, Hsin-Hsi, Opinion Mining and Relationship Discovery Using CopeOpi Opinion Analysis System. *J. Assoc. Inf. Sci. Technol.* **2009**, *60*, 1486–1503.
34. Wang, S. M.; Ku, L. W. ANTUSD: A Large Chinese Sentiment Dictionary. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia, 23–28 May 2016, pp. 2697–2702.
35. Azmi, A. M.; Altmami, N. I. An Abstractive Arabic Text Summarizer with User Controlled Granularity. *Inf. Process. Manag.* **2018**, *54*, 903–921, doi:10.1016/j.ipm.2018.06.002.