

The road to recovery: Sensing public opinion towards reopening measures with social media data in post-lockdown cities

Yiqiao Chen, Haifeng Niu^{*}, Elisabete A. Silva

Department of Land Economy, University of Cambridge, Cambridge, United Kingdom

ARTICLE INFO

Keywords:

COVID-19
Urban management
Public opinion
Social media data
Twitter
Recovery measures

ABSTRACT

The COVID-19 pandemic has resulted in cities implementing lockdown measures, causing unprecedented disruption (e.g. school/shop/office closures) to urban life often extending over months. With the spread of COVID-19 now being relatively contained, many cities have started to ease their lockdown restrictions by phases. Following the phased recovery strategy proposed by the UK government following the first national lockdown, this paper utilises Greater London as its case study, selecting three main reopening measures (i.e., schools, shops and hospitality reopening). This paper applies sentiment analysis and topic modelling to explore public opinions expressed via Twitter. Our findings reveal that public attention towards the reopening measures reached a peak before the date of policy implementation. The attitudes expressed in discussing reopening measures changed from negative to positive. Regarding the discussed topics related to reopening measures, we find that citizens are more sensitive to early-stage reopening than later ones. This study provides a time-sensitive approach for local authorities and city managers to rapidly sense public opinion using real-time social media data. Governments and policymakers can make use of the framework of sensing public opinion presented herein and utilise it in leading their post-lockdown cities into an adaptive, inclusive and smart recovery.

1. Introduction

Cities and other urban centres are where 56.2 % of the world population lives (UN-Habitat, 2020). Yet due to unprecedented restrictions on social gathering and movement, cities have been undoubtedly become the front lines in fighting the pandemic. After several months of lockdown, many UK cities are starting to ease their restrictions and move towards an early recovery. However, compared with the relatively blunt and straightforward measures taken to implement the lockdown, reopening cities is a more complex challenge. When and how to reopen cities are seen as a dilemma and trade-off between lives and economy. Policy makers must take into account the timing of the recovery process, the sectors to prioritise, the needs from different local communities, and, most importantly, the uncertainty of further outbreaks. Given these circumstances, governments tend to implement more adaptive measures, enabling them to gradually recover urban life by phases. For instance, in the UK, the recovery strategy published by the government emphasises that the process must proceed with ‘the utmost care in the next phase’ and highlights the requisite flexibility of the planned timeline for lifting restrictions (Cabinet Office, 2020). In the UK capital, the

London Transition Board and Recovery Board have been established to manage the complex process of opening up. It aims to reverse the social and economic loss, support communities, reduce inequalities and deliver a greener London, emphasising public engagement, involvement and inclusion during the recovery (GLA, 2020a).

In this process of slowly reopening cities, public opinion plays a vital role. On the one hand, the public response directly reflects inhabitant perceptions and expectations with regards to reopening measures, which in reality can be regarded as a form of collective behaviour (Blumer, 1948; Wlezien, 2017). On the other hand, it can be a valuable source of information which policymakers can utilise to adapt current measures and policies for the next phase. In the era of big data, social media has emerged as a major data source by which to sense public opinion, thus providing unique opportunities for supporting urban management from the bottom-up (Feezell, 2018; Hocht, Parycek, & Schollhammer, 2016).

This research aims to discover and ascertain public opinion towards the COVID-19 pandemic, the post-COVID-19 reopening measures and phases within a global health crisis context; resultantly, this paper provides critical and timely insights by which to understand and

^{*} Corresponding author at: Department of Land Economy, 19 Silver Street, Cambridge CB3 9EP, United Kingdom.

E-mail address: hn303@cam.ac.uk (H. Niu).

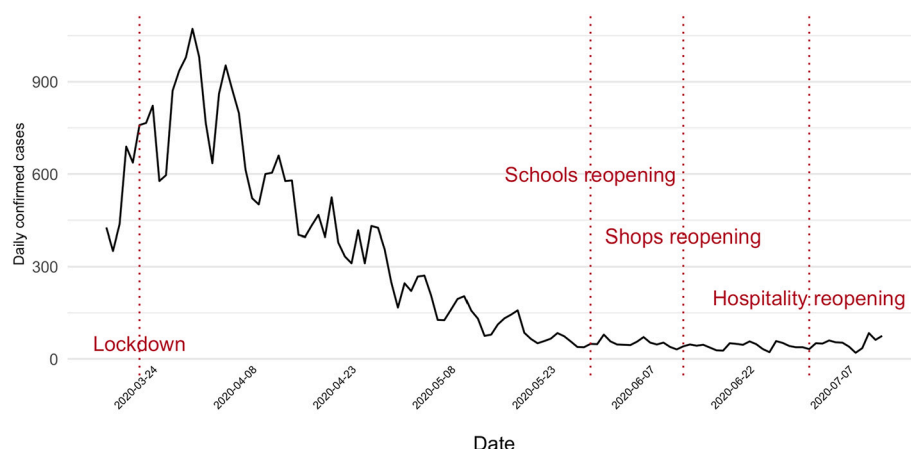


Fig. 1. Daily trends of COVID-19 confirmed cases and implementation of reopening measures in London (Data Source: Public Health England).

facilitate the post-pandemic recovery in London. The paper is organised as follows. Section 2 analyses a concise literature review on urban big data, social media and the COVID-19 pandemic. Section 3 explains the methodology employed herein, including data collection, data processing, sentiment analysis and topic modelling. Section 4 presents the results. Section 5 further discusses results, contributions, limitations and further directions. Section 6 summarises the key findings and concludes the paper.

2. Literature review

Government responses and urban management are crucial during the COVID-19. Main measures taken by governments include quarantine, self-isolation, digital surveillance, lockdown, and reopening (Jasiński, 2021; Lin, Lin, Yan, & Huang, 2021; Tan, Chiu-Shee, & Duarte, 2022). As these policy measures have significant impacts on people's daily lives, they draw great attention from the public and researchers. Managing COVID-19 is not only a public health issue, but also a political concern (Jasiński, 2021). Under this circumstance, since the beginning of the global pandemic, studies have been conducted on understanding policy measures in terms of effectiveness, consequences, stringency, acceptances, and opposition. For example, Zhang, Ji, Zheng, Ye, and Li (2020) evaluated the effectiveness of the lock-down strategy. Tan et al. (2022) explored the role of digital interventions such as digital quarantine measures during COVID-19. Guo, Chen, and Liu (2022) analysed the citizens' acceptance of health QR codes in Chinese cities. In the relevant literature, the most widely used data sources include social media, questionnaires, surveys, and administrative data (Guo et al., 2022; Kausshal & Mahajan, 2021; Lin et al., 2021; Liu et al., 2021; Yao, Yang, Liu, Keith, & Guan, 2021).

To better support urban management during emergent crises, social media is an important source. Social media data record human activities in cities with time, location, tags, texts, images and other profile information, allowing researchers to explore many aspects of urban management (Niu & Silva, 2020). Before the COVID-19 pandemic, real-time information from social media can support emergency management in public engagement, public-private communication, situation monitoring, social cohesion creation, intervention evaluation, and collaborative governance (Alexander, 2014; Gibbons, Nara, & Appleyard, 2018; Pereira, Parycek, Falco, & Kleinhans, 2018; Vayansky, Kumar, & Li, 2019; Xu et al., 2020). During the COVID-19 pandemic, social media data have been used in understanding government responses because it is one of the fastest ways to sense governmental communication strategies, political preference, public attention and sentiment (Ahmed, Rabin, & Chowdhury, 2020; Chen, Silva, & Reis, 2021; Shen et al., 2020; Tsao et al., 2021; Wu et al., 2022).

Although social media platforms provide real time and massive user

generated content, mining public responses to urban policies is still a significant challenge. First, social media data contain vast amounts of irrelevant information and how to select datasets for analyses of specific policy measures requires a well-designed process of data pre-processing. Second, posts on social media are textual and formed as unstructured data. Converting unstructured text of social media posts into insights requires the implementation of text mining techniques such as sentiment analysis and topic modelling. Sentiment analysis can be used to reveal the positive, negative or neutral tones of textual user-generated content (Nielsen, 2011). In order to explore public opinion towards specific issues/policies, researchers such as Chen, Silva, and Reis (2021) and Vayansky et al. (2019) conduct sentiment analyses to extract the sentiment polarity of the public. Topic modelling is another widely utilised approach in further investigating preeminent topics (Abd-Alrazaq, Alhuwail, Househ, Hamdi, & Shah, 2020). Topic modelling aims to extract main themes from unstructured documents (Isoaho, Gritsenko, & Makela, 2021). For example, Zhou, Tao, Rahman, and Zhang (2017) use topic models to identify hidden communities of tweets. Jiang, Qiang, and Lin (2016) assessed major concerns regarding a controversial infrastructure project through a topic modelling algorithm. Although these text mining methods have been separately used in previous studies, how to integrate those methods in sensing public responses to dynamically changing policies is still lacking. In this paper, we intend to propose a framework for monitoring public responses to policies from social media with the implementation of text mining methods.

In the current stage at the time of writing (i.e., recovery/reopen stage), many countries have begun to ease lockdown restrictions and are gradually reopening; thus, reopening strategies and discussions are the new focus in urban governance. COVID-19 has had a huge economic impact, resulting in many job losses; accordingly, many people want to return to work and are asking the government to reopen (Samuel, Rahman, Ali, Samuel, & Pelaez, 2020). On the contrary, as lockdown is an effective approach to reduce new cases and prevent dissemination of the virus, some people have criticised attempts reopening made during early June 2020 as risky and dangerous (O'Dowd, 2020). Faced with conflicting debates on reopening the country, it is necessary to explore and identify what the dominant sentiment is towards current or planned reopening strategies and whether the public is satisfied with these measures. Presently, comprehensive analyses of recovery strategies and phases are rare before and during the implementation of reopening policies, with less than 60 studies (published prior to July 2020) found on Web of Science. Among the small number of studies on reopening, researchers limit their analysis to reopening/recovery policies in different cases (predominantly in the US) with specific contextual focuses, such as schools and healthcare services (Brandenburg et al., 2020; Preskorn, 2020; Samuel et al., 2020; Sheikh, Sheikh, Sheikh, & Dhimi, 2020). In response to the limitations of the literature, this study seeks to

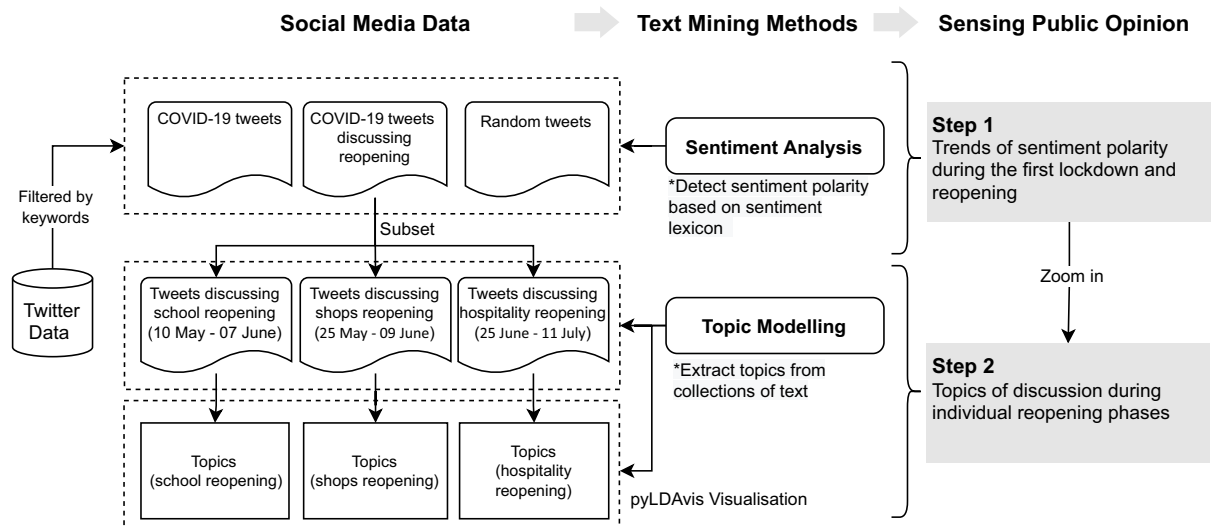


Fig. 2. Research framework of sensing public opinion regarding reopening measures from social media data.

understand up-to-date public perception via-a-vis three stages of reopening (i.e., schools, shops, and hospitality reopening) after the first national lockdown in Greater London.

3. Research design

3.1. Case study

Greater London initiated its lockdown officially on 23rd March 2020, being one of the worst affected regions in the UK. Initially, in March 2020, London was hardly hit by the coronavirus relative to other international cities; the daily confirmed cases increased sharply in March but gradually decreased in the following months after the city entered its lockdown phase (see Fig. 1). The effects of prolonged lockdown have during the past few months become more apparent, with people suffering from mental and physical health crises and with palpable anxiety in facing an increasingly uncertain world. When and how to recover the city naturally emerged as a hot debate for both government and the public, both on traditional media and social media. Many Twitter users in London have posted a great number of tweets discussing the pandemic, the reopening measures and other topics on a daily basis; this results in a vast and valuable dataset that contains insights regarding underlying public opinions. Concerning its relevant to policy, the Mayor of London and the Greater London Authority stated that public opinion should be valued, and all Londoners should be engaged in reimagining the new normal in the post-COVID-19 world, that is, during the recovery process (GLA, 2020b). Greater London is thus selected as the subject of analytical focus in this case study.

The lifting of lockdown restrictions in Greater London follows the recovery strategy published by the UK government. According to the so-called reopening roadmap, restrictions would be lifted step by step, depending on the spread and distribution of the virus. By the end of July 2020, the government announced a series of reopening measures to ease the first lockdown (see Fig. 1). As stated in the COVID-19 restriction summary offered by GLA,¹ the key reopening policies are:

- i. From 1st June, schools start to open for more children²;

- ii. From 15th June, non-essential retail stores are allowed to open³;
- iii. From 4th July, more non-essential business and services such as personal care, hospitality, public places and leisure facilities are allowed to open.⁴

The news/announcements of reopening measures were released around 20 days before the actual changes were to be implemented. On 10th May, the government stated that 1st June would be the earliest possible date for reopening schools. On 25th May, the date when non-essential shops could reopen was declared. Then on 23rd June, a set of lockdown-easing measures (to come into effect from 4th July) were reported on the news.

These reopening measures essentially delineate the road to the recovery in the UK, including London. In this research, we define three chronological phases, namely the **schools reopening**, **shops reopening**, and **hospitality reopening** phases. We then investigate the three phases of recovery in London using the three measures above.

3.2. Research framework

To sense the public opinion regarding the reopening measures during the first lockdown and recovery stages, we use Twitter data as the main data source because people use this media to disseminate the information and express opinions towards the reopening measures in real-time. A two-step approach is developed as follows (Fig. 2).

- i. Detecting the trends of sentiment polarity during the first lockdown and reopening.

This step first measures the daily trends of sentiment polarity (i.e., positive, negative and neutral) to detect the general sentiment in London per day during the first lockdown and reopening. To detect the sentiment changes impacted by the reopening measures, three sets of Twitter data are selected. They are 1) COVID-19 related tweets; 2) COVID-19 related tweets discussing reopening measures; and 3) random tweets.

- ii. Extracting topics from discussions about individual reopening measures.

To further understand the public responses to individual reopening

¹ <https://data.london.gov.uk/dataset/COVID-19-restrictions-timeseries>.

² <https://www.gov.uk/government/news/pm-confirms-schools-colleges-and-nurseries-on-track-to-begin-phased-reopening>.

³ <https://www.gov.uk/government/speeches/business-secretarys-statement-on-coronavirus-19-9-june-2020>.

⁴ <https://www.gov.uk/government/news/pm-announces-easing-of-lockdown-restrictions-23-june-2020>.

Table 1

Total number of tweets used in the study.

Month	Number of random tweets	Number of COVID-19 tweets	Number of reopening tweets	Number of school reopening tweets	Number of shop reopening tweets	Number of hospitality reopening tweets
March 2020	1,480,297	–	–	–	–	–
April 2020	1,209,415	3,726,928	12,422	2325	2359	1372
May 2020	1,243,531	3,649,066	31,258	10,476	4881	3563
June 2020	1,153,457	2,168,863	15,914	2911	1720	5375
July 2020 ^a	–	10,003	84	2	4	54

^a Until 15th July, inclusive.

measures, we subset the dataset of COVID-19 related tweets discussing reopening measures. Three subsets of tweets are filtered by time periods and relevant keywords (e.g., shops, stores, schools, cafes, and bars). By applying topic modelling techniques, we extract specific topics from each subset of reopening tweets and visualise the result in an interactive manner.

3.2.1. Twitter data collection and pre-processing

The social media data analysed in this investigation were collected through Twitter's developer API (<https://developer.twitter.com/en/docs>), using R package *rtweet* (Kearney, 2020). With the parameters of the case study set, we collected relevant tweets posted in Greater London from 1st March 2020 to 15th July 2020. We conducted a geo search by setting the given radius at 20 miles, with the point of origin set as the central coordinates of London, enabling us to cover the administrative boundary of Greater London. The geo searches first find all tweets which are located within the geocode and then filter these tweets (geotagged tweets and non-geotagged tweets) created by users located within the defined search area. English was also set to limit the Twitter feed languages. The collected tweets include geotagged and non-geotagged posts. The first group contains random tweets in London, where keywords have not been pre-determined during the data collection process. The second group are COVID-19 related tweets, with keywords 'COVID19 OR coronavirus* OR COVID'.

For each group, we first manually removed tweets published by bot accounts and news accounts (which could be identified based on the high frequency of posts and self-declared account descriptions). We calculated the dominant users who tweeted posts frequently and further analysed the top 50 users' accounts by frequency. By reviewing the users' descriptions and other profile information (e.g., username and profile image), we removed bot accounts and news accounts from the dataset. Having finished the user screening, we extract the reopen-related tweets from all COVID-19 tweets by filtering the keyword 'reopen*'. In order to analyse three phases of reopening, three groups of reopening tweets were subset from the reopen dataset. We applied the keyword 'school*' to filter school reopen tweets. For the second phase, keywords of 'shop*' and 'store*' are used to filter the dataset. The key discussions on London's reopening on 4th July 2020 were bars, pubs, restaurants, hairdressers, and business; accordingly, keywords of 'pub*', 'restaurant*', 'hairdresser*', 'bar*', 'business*', 'café*', 'barbershop*' were set to filter and collate the relevant tweets. The numbers of tweets in each data group are presented in Table 1. We further cleaned unstructured tweets by removing URLs, mentions, noisy words (i.e., Re-Tweets), newlines and extra whitespaces. Then, stop words such as 'the', 'that', and 'on' were removed using Python package *NLTK*, a natural language toolkit to remove useless words (see Loper & Bird, 2002). Finally, *WordNet Lemmatizer* from the package *NLTK* was utilised to covert the words into their base forms.

3.2.2. Text mining methods

3.2.2.1. Sentiment analysis. Sentiment analysis is a popular approach by which to explore insights from social media data. After cleaning all the collected tweets, we conducted sentiment analysis through via the Python package *AFINN* 0.1 (Nielsen, 2011). This is one of the fastest and most commonly used sentiment analysis tools and has been broadly applied. For instance, previous urban studies applied this method in exploring public sentiment towards urban phenomenon and urban planning measures (Silva, Liu, Kwon, Niu, & Chen, 2020; Hollander & Renski, 2017). *AFINN* sentiment analysis is a lexicon-based method to score each word by comparing it to scores of an existing English word list (Al-Shabi, 2020). After scoring each word in the tweets, the result of a post sums up all scores in a sentence. Each word has a score between −5 to 5. A negative score denotes a negative sentiment while a positive result denotes a positive sentiment, with zero denoting neutral sentiment (Nielsen, 2011). After generating the scores for all tweets, public attitudes and emotions can be revealed. For all groups (see Table 1), we calculated the means of everyday tweets to provide daily sentiment trends. The daily means of random tweets show the general sentiment in London per day during the analysed period. The sentiment results in the COVID-19 group and reopening group thus reveal public feelings, emotions and sentiment towards the pandemic lockdown and related reopening measures.

3.2.2.2. Topic modelling. Topic modelling is one of the most powerful text mining tools for exploring semantic structure from a collection of texts. Latent Dirichlet Allocation (LDA), as a generative probabilistic model, is commonly employed to extract topics from a collection of documents (Capela & Ramirez-Marquez, 2019; Taecharungroj & Mathayomchan, 2020). In the LDA topic model, the collection of documents is referred as the *corpus*; items within the corpus are referred to as the *document*, with specific words in documents called *terms*. The LDA model assumes that a document is generated according to the following process: 1) decide the number of words N that the document will include prior to randomly choosing a distribution over topics; 2) generate each word in the document. In step two, the model probabilistically draws one of the K topics according to the distribution over topics sampled above, and probabilistically draws one of words according to the topic's multinomial distribution. Based on this generative model, the LDA model backtracks from the documents (as defined herein) to discover the topics that are likely to have produced the corpus. In essence, the assumption of the LDA topic model is that each document in the corpus is a mixtures of K topics that are characterised by terms with certain probabilities. Each latent topic is characterised by a distribution over a fixed vocabulary. The details of LDA can be found in Blei, Ng, and Jordan (2003).

In this study, LDA topic modelling is utilised so as to reveal the latent topics inherent to the data, with probabilities of terms from the documents, which are tweets discussing the reopening measures. To further explore topics embedded within captured social media discussions, we

Table 2

Corpora during individual reopening phases for topic modelling.

Corpora	Time range	Number of Tweets	Number of terms
Corpus (school reopening)	10 May 2020–07 June 2020	18,842	3636
Corpus (shop reopening)	25 May 2020–09 June	17,459	4500
Corpus (hospitality reopening)	25 June 2020–11 July	3350	1700

subset the reopening tweets for each measure by setting the time range to commence from the initial announcement (10th May, 25th May, 23rd June 2020, respectively) to the week after the measure was implemented (see time range in Table 2). This is because the discussions normally start from the release of government reopening measures and continued after their actual implementation. Thereafter, we conduct topic modelling separately for each reopening measure corpus with the LDA model by using the Python package *Gensim* (Rehurek & Sojka, 2010). In the process of training the LDA model, we first created a dictionary representation for all tweets. Additionally, we removed rare terms and common terms based on their term-document frequency.

Then, we transformed the tweets into a vectorised form with the bag-of-words representation. These vectorised tweets were input for LDA training. The output was a list of topics with probabilities ascribed to each topic. We computed the LDA model for different values of k , ranging from 2 to 10. This is because we intend to extract at least two topics while concurrently limiting the number of topics so as to avoid generating too many similar topics. For each reopening measure corpus, we compared inter-topic distance maps for different values of k in the LDA model and selected the optimal k which returns only distinct topics (i.e., less overlapping of topics on the inter-topic distance map). We operate this process by utilising *pyLDAvis*, a Python library developed by Sievert and Shirley (2014), to compute the inter-topic distances based on Jensen-Shannon divergence and to visualise the inter-topic differences into interactive maps.

4. Results

We conducted descriptive analysis, sentiment analysis, and topic modelling according to the processes explained in Section 3, in order to ascertain and explore public attention, sentiment and main discussions evidenced by the collected data. The results of overall attention towards the coronavirus, the reopening measures and three different phases are

presented first. The daily means of sentiment scores of random tweets, COVID-19 discussions, and recovery measures are introduced. Narrowing our focus to the three different opening up phases, we demonstrate the topic modelling results of discussions on the recovery measures in a specific timeframe in detail.

4.1. Public attention

London exhibits a particularly active Twitter community, which have posted a substantial number of tweets (see Fig. 2). In April, Londoners posted around 155,000 tweets related to COVID-19 per day, while the average number of posts per day in May was about 120,000. The following months witnessed a sharp decline in COVID-19 discussions. The daily frequency of tweets was 72,000 in June and less than 700 in July. Public attention towards the pandemic appears to be decreasing as time goes by. This phenomenon could be explained by the fact that the situation in London was improving, since the number of daily confirmed cases continued to decrease over time. It might also be explained by the topic becoming less popular as people gradually got used to the coronavirus and thus became less interested in discussing it. Other social-political events have attracted more attention from Londoners. For example, the death of George Floyd and the Black Lives Matter protests became new and emergent hot topics on Twitter since the end of May. During the lockdown period, the daily discussion on COVID-19 was more than 100,000, being one of the most recurring trending topics during the past four months (March to July 2020).

For Twitter discussions related to reopening the city, we focus our investigation on three distinct phases, as explained in Section 3.1. For each phase, we identify the date that the government released the related announcement and the actual reopening date (see the labels in Fig. 3). Generally, the topic of reopening/recovery comprised less than 5 % of posts in all COVID-19 related tweets, as illustrated in Fig. 3. The daily frequency of reopening posts was 670 before London's re-opening on 4th July 2020. Following the government's announcement of the news, people increasingly discussed the reopening measures. However, when it comes to the actual date of implementation, discussions focused on the relevant topics did not reach another peak.

We further investigate the tweets related to different phases of reopening the city. In the first phase, the government announced that the earliest date for reopening schools (for some students only) was 1st June 2020. The number of schools reopen discussions is illustrated by the blue line in Fig. 4. Both the date when the government released the news, and the reopening data itself, did not witness many discussions. The 1st June date did not exhibit a spike in debate, probably because

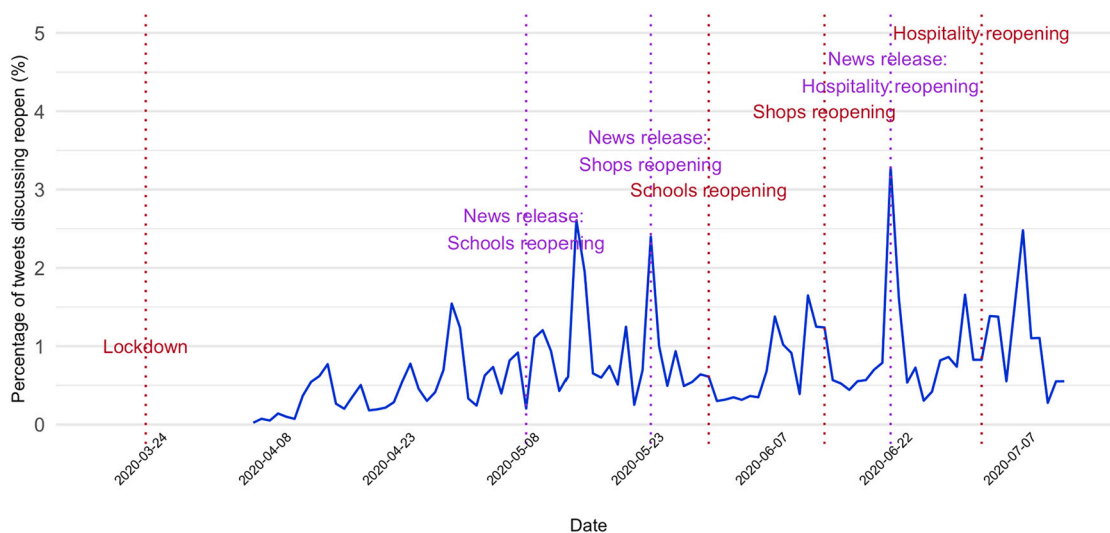


Fig. 3. The proportion of reopening discussions under COVID-19 discussions in London (Data source: Twitter Streaming API).

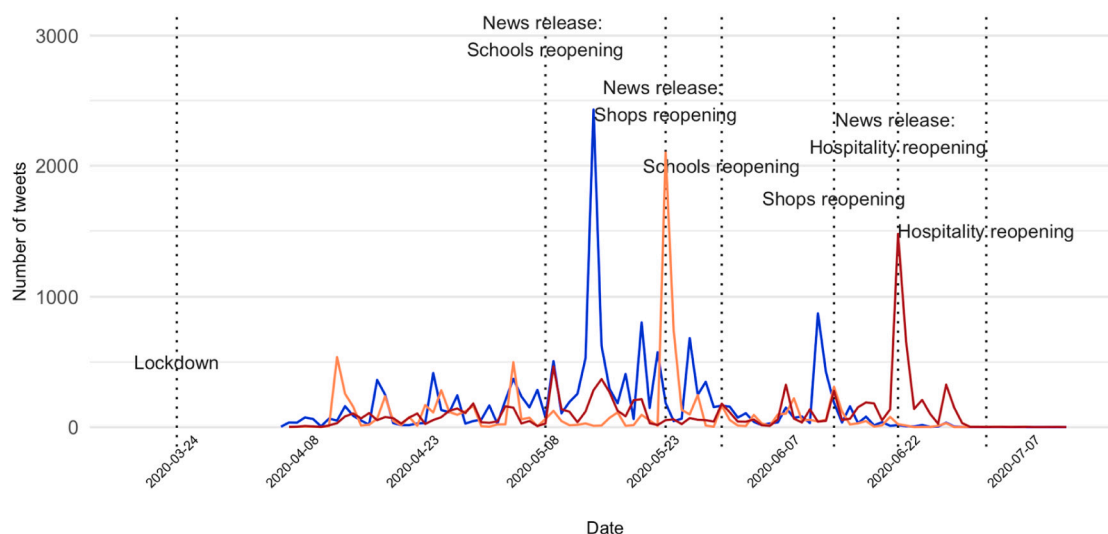


Fig. 4. Daily trends of three reopening measures related tweets in London (Data source: Twitter Streaming API).

Note: blue line – schools reopening, orange line – shops reopening, red line – hospitality reopening. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

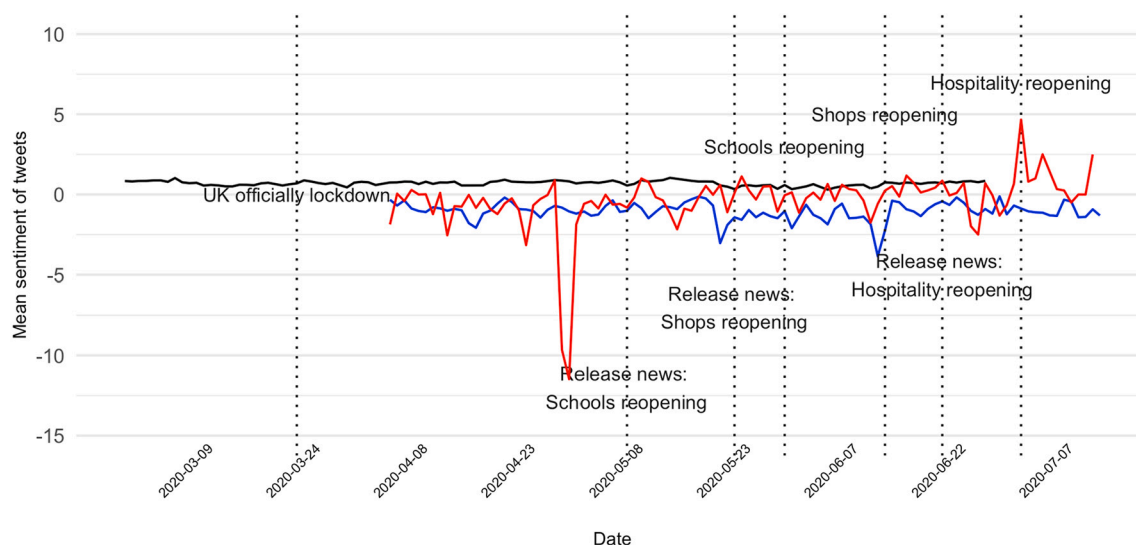


Fig. 5. Overview of the sentiment of tweets in London (Data source: Twitter Streaming API).

Note: blue line – COVID_19 tweets, red line – reopening tweets, black line – random tweets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

only a small proportion of students were affected. In terms of the announcement date on 10th May, less than 0.05 % of COVID-19 related tweets mentioned schools reopening. This may be because people in London had a lower awareness of the government's guidelines; in support of this view, according to a survey conducted by the ONS, Londoners stated that they do not have enough information about the UK's plan (ONS, 2020). The week following the announcement saw a growing number of tweets. The Twitter discussion reached its peak on 16th May, a week after the announcement. Debate intensified, especially regarding government guidance of reopening schools and what to do if a child shows any symptoms. The Mayor of Liverpool stated that Liverpool's schools would ignore the government's requirement of reopening on 1st June. Although this issue occurred in Liverpool, Londoners were evidently impressed by the news and expressed their worries and concerns about the school reopening, illustrated by the rise in tweet volumes.

During the second and third phases, the discussion reached its high

point on the dates that the news (of reopening measures) was released (see orange and red lines in Fig. 4). Similar to the school reopening context, the actual date of non-essential shops reopening (15th June) and restaurants/pubs/businesses reopening (4th July) did not attract much public attention on the social media platform.

4.2. Public sentiment

In general, random tweets posted by Londoners are predominantly neutral and slightly positive from March to June 2020, ranging between 0.30 and 1.33 (illustrated by the black line in Fig. 5). The daily means of random topics remain stable, with the standard deviation in the range of 2.31 to 3.22. The average sentiment of discussions on Twitter did not change dramatically during the lockdown or the respective recovery stages.

The public sentiment towards COVID-19 related topics was negative and consistently below the general sentiment in the city, as illustrated by

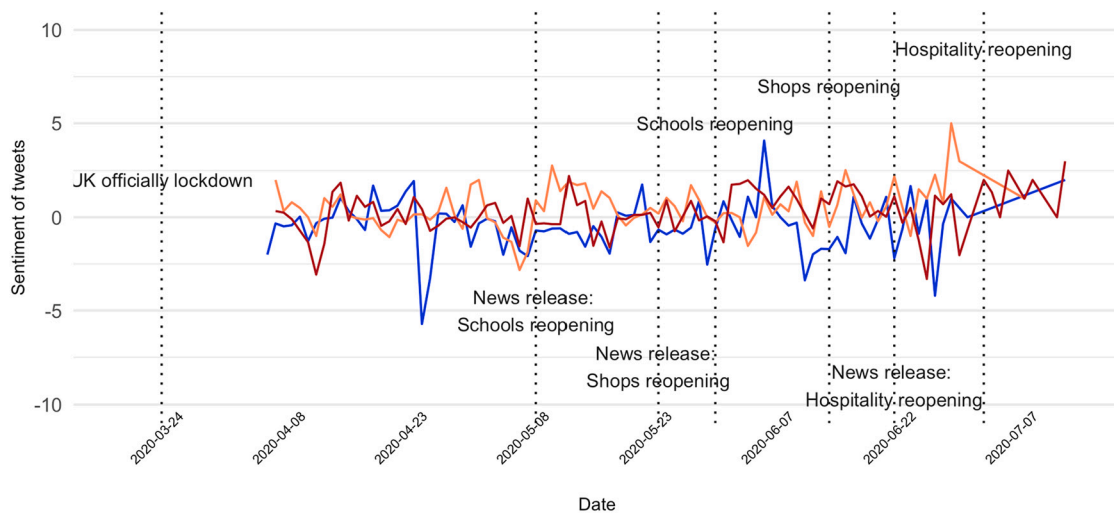


Fig. 6. Daily trends of discussions on three reopening measures in London (Data source: Twitter Streaming API).

Note: blue line – schools reopening, orange line – shops reopening, red line – hospitality reopening. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the blue line in Fig. 5. The highest daily mean of COVID-19 posts was -0.12 on 1st July, while the worst mean was -3.9 on 14th June. The emotions of people when discussing COVID-19 related topics are evidentially more variable, indicated by the higher standard deviation (2.95 to 5.54) in sentiment results. We observe that public sentiment fluctuated significantly when expressing negative ideas. The lowest mean and the highest standard deviation occurred on the same day, when people were complaining that the UK had the worst coronavirus death rate in Europe and expressed concerns about the outbreak risks and anti-lockdown protests.

The daily average sentiment scores of reopening issues (illustrated by

the red line in Fig. 5) were mostly between the scores of the general sentiment and COVID-19 related sentiment, that is to say, somewhat negative. The daily sentiment towards all reopening varies between -5.71 and 5.00 . The average sentiment is 0.12 , which is lower than the that of random tweets (0.70). When discussing the reopening measures, such as timelines, preparations, and risks, the data reveals that people were predominantly worried in April and May; however, sentiment scores rose moderately during June and July. The increasing sentiment results indicates that Londoners are becoming more positive an optimistic towards the recovery process. The increasing trend in sentiment may be due to the much-welcomed news that confirmed cases were

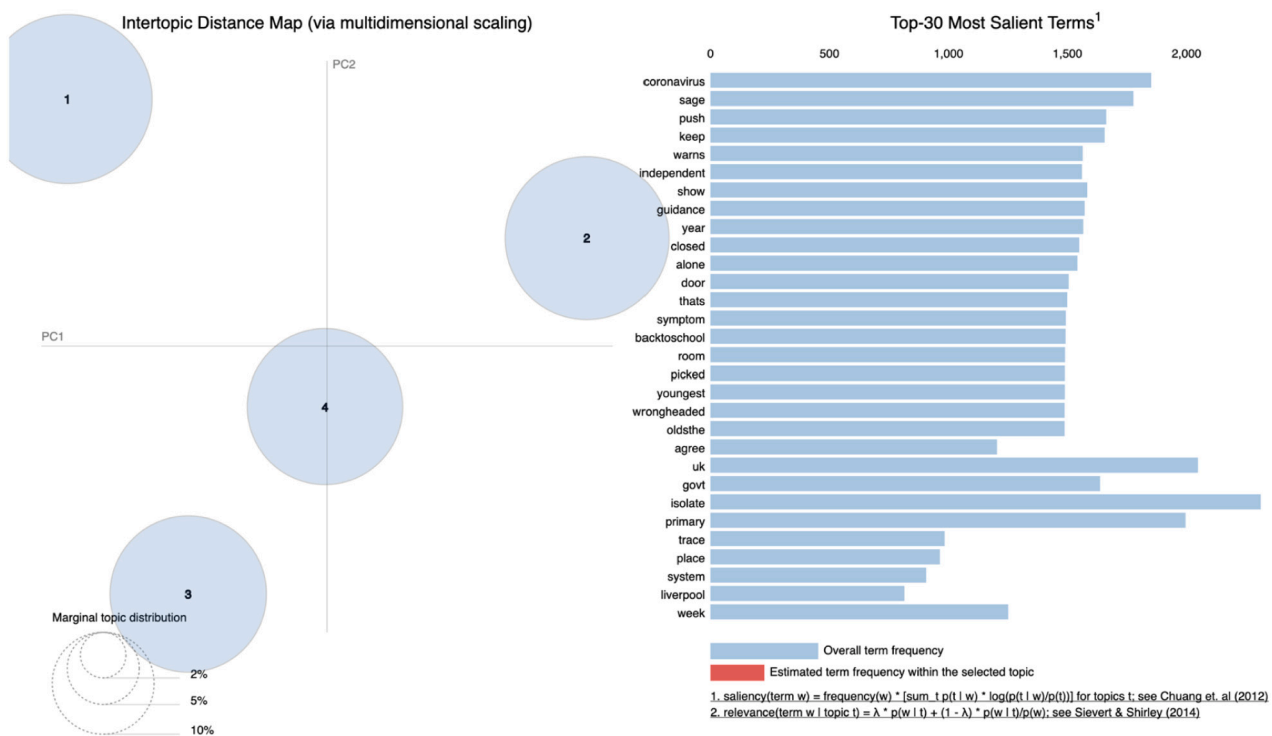


Fig. 7. Four topics and terms frequency for school reopening discussion found in tweets.

On the left panel, each circle refers to a topic; circle sizes denote the importance of the topic over the corpus (i.e., the larger the circle, the more pertinent the discussion). The distance between circles indicates the similarity between topics. On the right panel, a horizontal bar chart lists the top 30 most relevant terms with the overall term frequency.

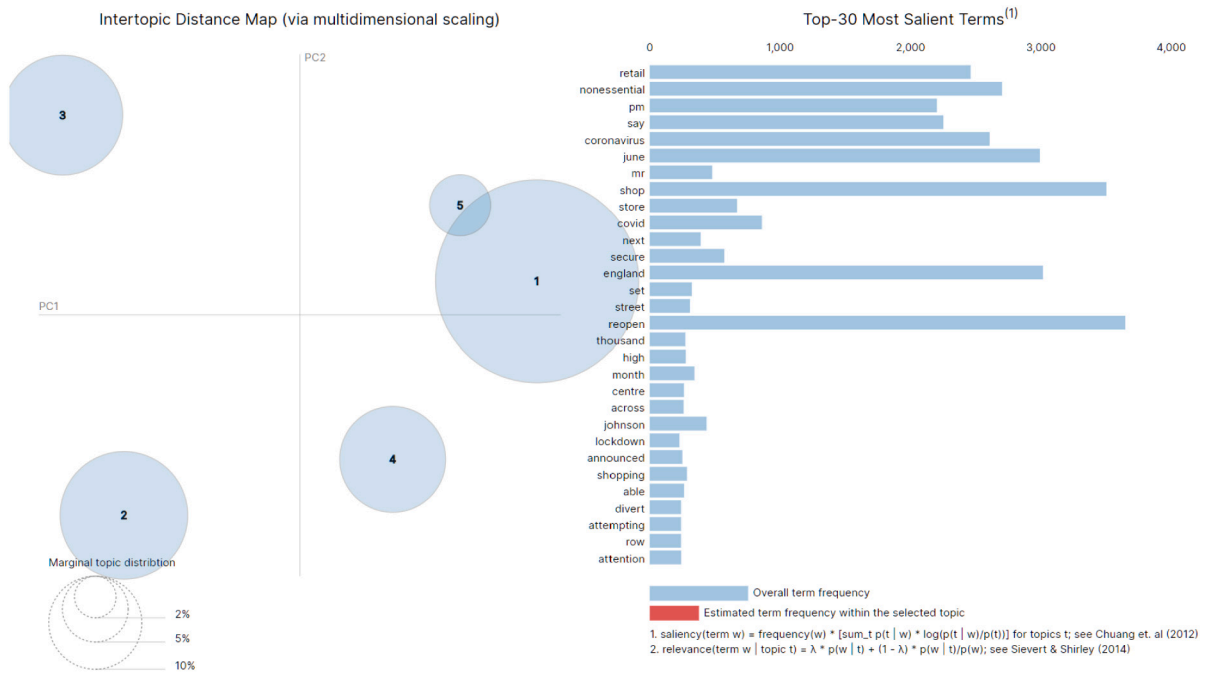


Fig. 8. Five topics and terms frequency for shop reopening discussion found in tweets.

On the left panel, each circle refers to a topic; circle size denotes the importance of the topic over the corpus. The distance between circles indicates the similarity between topics. On the right panel, a horizontal bar chart lists the top 30 most relevant terms with the overall term frequency.

decreasing; concomitantly, people were more comfortable to venture out, being tired of staying at home. Very positive sentiment could be identified on the 4th of July, with a very limited number of tweets collected on that day. The most negative sentiment is evident on 2nd May, when people were clearly upset by news such as the arrest of a suspect accused of murdering an NHS worker, Italy's struggle to recover, and questioning of the fraudulent inflation of COVID-19 statistics.

In order to understand public opinion towards different reopening measures, we assess the temporal sentiment of discussions towards schools reopening, shops reopening, and hospitality reopening, respectively (Fig. 6). Opinions regarding schools reopening was slightly negative than the other reopening issues. The sentiment scores of tweets discussing schools reopening ranged between −5.71 and 4.10. Londoners tweeted about their concerns about students' safety if they need

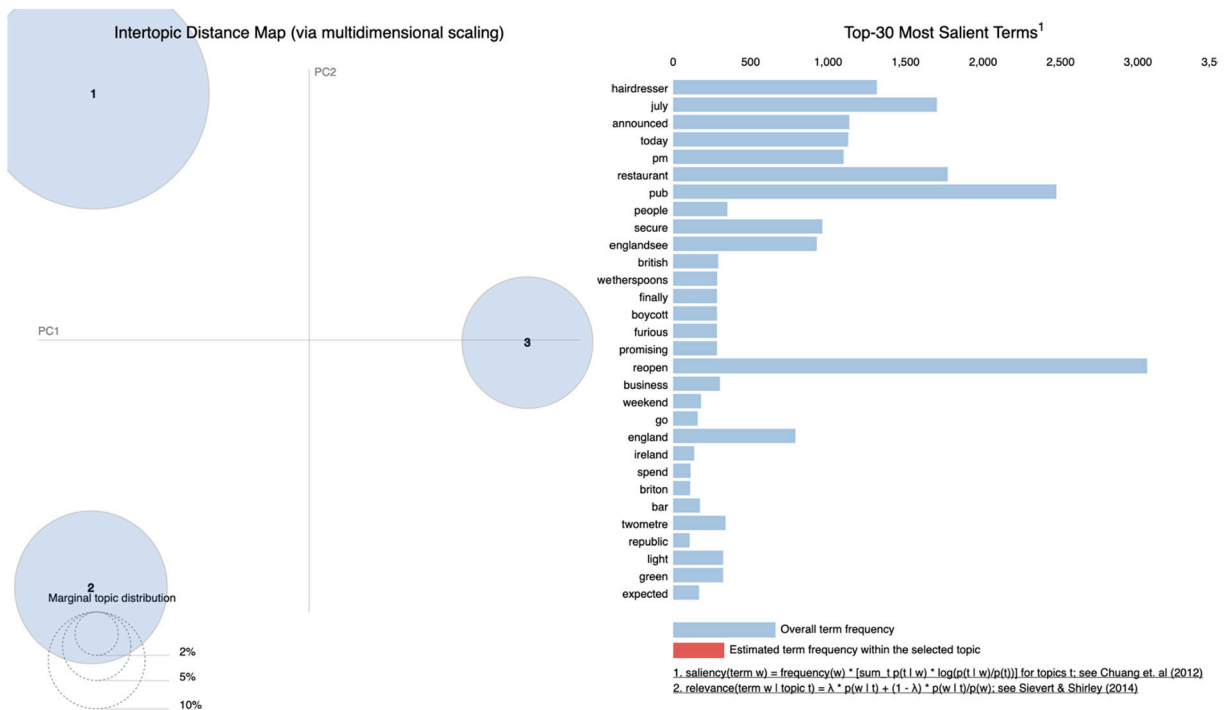


Fig. 9. Three topics and terms frequency for hospitality reopening discussion found in tweets. On the left panel, each circle refers to a topic; circle size denotes the importance of the topic over the corpus. The distance between circles indicates the similarity between topics. On the right panel, a horizontal bar chart lists the top 30 most relevant terms with the overall term frequency.

Table 3

Terms within topics extracted from school reopening discussion.

	Topic 1	Topic 2	Topic 3	Topic 4
Topic	Push , keep, show, guidance, year,	SAGE , warns , independent, week,	Agree, say, COVID, Liverpool , pupil,	Coronavirus, lockdown, June, risk,
Policy	closed, alone , door, symptom, back to	trace, place, England, test, system,	government, Monday, u, open,	return , scientist, case , new, infection,
1	school , room, picked, youngest , olds	safe, first, sir, chaired, David ,	please, least, tell, last, sept, question ,	Michael , Gove , live, warn, Johnson ,
Corpus	the, wrongheaded , child, primary ,	functioning, king, need, early, June,	sec, ed., death, lab, none, shadow ,	England , spreading, help, stage ,
	govt, UK, isolate, COVID, story, made,	next, isolate, government, rate ,	leader , day, must, Joe , mayor	minister, fight, Boris , step, nursery,
	flat, right, children , fall, association	another, set, advice , change,		plan, people, many, weve, reduce,
		exclusive, information, headteachers		substantially, series

to attend schools, concerns which often took a negative tone. Positive sentiment can be observed when the government announced extra support in reopening the schools.

The sentiment towards reopening non-essential shops fluctuates around 0, varying between −2.82 and 5.00. The discussion on reopening pubs, restaurants and other such businesses on 4th July exhibits slightly more stable emotions than on reopening schools and shops, with average means ranging between −3.30 and 3.00.

4.3. Main topics for reopening discussions

To explore more detailed insights in reopening discussions, we further investigate the latent topics across three reopening measures with the results from the LDA model and its illustrative web-based visualization tool (pyLDAvis). In the figures below, circles refer to the topics extracted from the set of tweets and the size of circles indicates the proportions of the topics in the corpus (i.e., the importance of the topic over the whole corpus). The right-hand side panel lists the most relevant terms with the overall term frequency. As mentioned in Section 3.2.2.2, we choose the optimal number of topics for each corpus by comparing the inter-topics distance maps with different number of topics.

4.3.1. Schools reopening

According to Fig. 7, there are four separated topics identified in school reopening tweets. Topic 4 covers the statement from the government, including the public talk regarding the possibility of reopening schools by Michael Gove (the UK Minister for the Cabinet Office) on TV and the official statement by Prime Minister Boris Johnson. The remaining topics essentially echo what we found in sentiment polarity for school reopening in the previous section, which is an overall negative attitude on the school reopening. Topic 1 refers to instances of negative feedback regarding school reopening measures, with terms such as ‘push’ and ‘wrongheaded’. Moreover, this topic focuses on the group affected by this policy by mentioning terms such as ‘primary school’, ‘youngest’, and ‘children’ and so on. Topic 2 is related to warnings and concerns expressed by the Scientific Advisory Group for Emergencies (SAGE) with terms such as ‘SAGE’, ‘scientist’, ‘David’ (the Chair of SAGE), ‘warns’ and ‘early’. Topic 3 reveals the opposition inside the UK government, such as the concern made by Liverpool Mayor Joe Anderson and opposition expressed by the shadow cabinet.

4.3.2. Shops reopening

Five topics are extracted from the shops reopening tweets (Fig. 8). As shown in Table 4, Topic 1 gathers a group of terms regarding the announcement of reopening shops, especially highlighting the thousands of High Street shops and the opening of city centres. Similarly, Topic 5 also relates to the release of the measure, discussing the change in the government slogan from ‘Stay at home’ to ‘Stay alert’. Topic 2 reflects the feedback from the market and the projected confidence in providing people protection, indicated by terms such as ‘providing’, ‘able’, and ‘secure’. In Topic 3, a series of positive terms such as ‘great’, ‘happy’ and ‘wow’ were expressed in public discussion, indicating the public welcomes the reopening of non-essential shops. However, Topic 4 contains terms such as ‘chaos’, ‘social distancing’ and ‘safety’, which is illustrative of concerns regarding shops reopening, especially with

regards to obeying social distancing rules inside shops.

4.3.3. Hospitality reopening

As illustrated in Fig. 9, three topics are extracted from the hospitality reopening tweets. Topic 1 is associated with reopening measures that apply to different business and services. The terms ‘restaurant’, ‘pub’ and ‘hairdresser’ were widely discussed, while other services such as ‘cinema’ and ‘hotel’ were less talked about (see Table 5). Topic 2, with terms such as ‘business’ and ‘economy’ and ‘back’ reflects the connection among the public about how the new reopening measure is expected to lead to economic recovery. Negative discussions regarding hospitality reopening measures were fewer than the previous two, although Topic 3 specifically captures the ‘boycott Wetherspoons’ event, which occurred during the reopening of pubs.

5. Discussion

5.1. Validation of results

To validate our results, we compared our findings with an official survey - the London COVID-19 online diary.⁵ The opinion research team at GLA organised a COVID-19 online diary to capture the opinion of 20 citizens in London during the COVID-19 pandemic. The online survey started in mid-May 2020 and ended in mid-July 2020, which overlaps with the time period of this present study.

Londoners experienced wide-ranging emotions during the lockdown, as evidenced in the report (GLA Opinion Research, 2020b). The sentiment scores relate to COVID-19 Twitter discussions analysed herein also show a wide range of sentiment (between −0.12 and −3.9), with a high standard deviation (from 2.95 to 5.54) in sentiment results. Moreover, the report states that the main emotions Londoners felt were that of being fearful, anxious, and frustrated (GLA Opinion Research, 2020b). The sentiment results concur with these findings, with this study also indicative of similarly negative emotions.

Regarding reopening policies, the summary of the official survey echoes our results in Sections 4.2 and 4.3. Respondents mention perceived risk and education attainment when discussing schools reopening (GLA Opinion Research, 2020a). Topic 1 and 2, in discussions of schools reopening, reveals concerns about associated health risks (see Table 3). For reopening non-essential shops, Londoners showed ‘general support for re-opening non-essential shops’ (GLA Opinion Research, 2020b). The relatively positive sentiment towards reopening shops relative to reopening schools, along with Topic 3 in discussing shops reopening, both suggest a general groundswell of support for shops reopening (see Table 4). When discussing hospitality reopening, Londoners were more prone to question ‘how the government have handled it’ (GLA Opinion Research, 2020b). Similarly, our findings reveal that the public questioned how hospitality reopening could lead to an economic recovery (see Topic 2 in Table 5). Overall, our results are concordant with the findings from the London COVID-19 online diary, supporting the validity of our findings.

⁵ <https://data.london.gov.uk/dataset/COVID-19-online-diary>.

Table 4
Terms within topics extracted from shop reopening discussions.

Topic	Topics			
	Topic 1	Topic 2	Topic 3	Topic 4
Policy	Mr, store, reopen, Johnson, next, set, month, shop, COVID, England, thousand, high, street, shopping, secure, centre, across, department, announced, attention	COVID, secure, able, reopen, shop, store, market, car, outdoor, showroom, June, extended, providing, new, week, case, guideline, open, like	June, England, nonessential, reopen, shop, retail, say, coronavirus, pm, great, happy, map, route, wow, Boris Johnson, news	Reopen, shop, coronavirus, plan, rule, Ireland, social, June, distancing, lockdown, England, chaos, eased, travel, republic, nonessential, measure, John, Lewis, safety
Corpus				Shop, meet, stayalert, nonessential, lockdown, government, guideline, coronavirus, reopening, next, week, retail, reopen, ease, start, news, street, given, announcement, green

5.2. Contributions, limitations and implications

5.2.1. Academic contributions

This paper has addressed the following objectives and filled in the related research gaps. First, we contribute to existing studies by exploring public responses to government measures and focusing on specific policies in the recovery stage. During the pandemic, many researchers analysed government responses from the government side and mainly used data from official accounts (Górska, Dobija, Grossi, & Staniszweska, 2021; Tsao et al., 2021). We particularly explore government measures from the public side, which supports public involvement and collaborative governance. Additionally, for sensing public attitudes through social media during the pandemic, a large number of studies focus on lockdown and digital tracking measures (Guo et al., 2022; Lin et al., 2021; Zhang et al., 2020). According to review by Tsao et al. (2021), only one out of 81 studies has considered the recovery stage. This study timely contributes to the existing literature by investigating public opinion in the recovery stage. Considering the empirical contribution, most of the studies related to reopening considered the regions/cities in the US (e.g. Brandenburg et al., 2020; Kaufman, Whitaker, Lederer, Lewis, & McClellan, 2020; Preskorn, 2020; Vest, Blackburn, & Yeager, 2021). To provide knowledge of reopening UK cities, we conducted an empirical study encompassing Greater London, which has hitherto not been investigated.

Secondly, this paper developed a two-step framework of monitoring public responses to reopening policies from social media with the implementation of text mining methods. This framework takes full advantage of the collective and dynamic nature of social media which provides a vast amount of public discussion. Moreover, this framework incorporates text mining methods supported by natural language processing and machine learning techniques to extract timely insights from large-scale and unstructured user-generated content. Compared with the previous studies that explored government responses towards the pandemic predominantly comprised descriptive analyses and traditional content analysis (e.g. Rufai & Bunce, 2020; Sutton, Renshaw, & Butts, 2020), this study advances analytical methods by utilising sentiment analysis and topic modelling in supporting government responses during crises.

5.2.2. Policy/practical implications

Distinct from this paper's academic contributions, this novel research also generates practical implications on urban policy making. In London, the announcements and news items covering reopening measures attracted attention from the public. This attention – and subsequent online discussion – enables the government to sense collective sentiment prior to the actual implementation of measures. For sentiment-sensitive measures such as reopening schools, public opinion can be monitored in quasi-real-time, whereby government can redesign or redirect measures according to public reactions towards the policy. If there should be a third lockdown and subsequent reopening measures put in place, governments can track public attitudes in real-time through social media, and thereby ascertain the public's main concerns.

Future studies can explore public attitudes towards the latter lockdown and reopening measures following the resurgence of COVID-19 cases which commenced shortly after the reopening of cities. London and many other UK cities have been put into second or third lockdown. These practical implications can provide insights to support or alert governments prior to the actual implementation of policies, thus minimising the risk of exacerbating public dissatisfaction. This allows governments to make responsive urban policies during periods of extreme uncertainty. In addition to reopening measures, further investigations can extend our application to other government actions, such as economic recovery plans and vaccine measures. Moreover, researchers can conduct comparative studies into public opinion towards the COVID-19 pandemic and related government responses in different areas/cases, so as to illustrate the heterogeneity (or lack thereof) of public reactions

Table 5

Terms within topics extracted from hospitality reopening discussion.

	Topic 1	Topic 2	Topic 3
Topic Policy 3 Corpus	Reopen, restaurant , July, pub , hairdresser , announced, today, COVID, pm, secure, England, coronavirus, barber , twometre , rule, light, green, reduced, given, COVIDsecure, able, provided, following, guidelinesstayalert, Johnson, Boris, cinema , plan, hotel	Business , reopen, bar, COVID, lockdown, new, case , coronavirus, measure , government, pub, guidance, back , safely , read, many, make, spike, economy , safety, safe, day, sure, see, reopened, guideline, remain, thing, expected, ordered	Reopen, pub, people, British, Wetherspoons , finally, promising, furious , boycott , weekend , go, Ireland, spend, Briton, republic, expected, ready, England, get, help, know, good, key, register, link, preparing, businessshow, webinars, open

worldwide.

In addition to understanding issues related to COVID-19, the approach we employ in this study can be easily transferred to other cases/cities and other issues, such as future crises involving extreme weather, earthquakes, disease pandemics, and economic fluctuations. The text-mining method employed in this study can extract insights from social media data, newspaper comments, and other documents and thereby reveal predominant or prevalent emotions, topics, and preferences towards specific issues, which can help researchers and policy-makers to comprehensively understand critical matters and make better decisions.

5.2.3. Limitations and further directions

This research is not without its limitations. Primarily, limitations can be identified in data sources and methods. In terms of data sources, social media data have been criticised due to its representativeness problem (see [Martin, Julian, & Cos-Gayon, 2019](#); [Niu & Silva, 2020](#)). For instance, at the global level, there are only 1 % of Twitter users that contribute to geotagged Tweets, and only 1 % of all Tweets can be accessed via Twitter Streaming API ([Morstatter, Pfeffer, Liu, & Carley, 2013](#)). On top of this, social media users significantly under-represent certain groups, such as the elderly and ethnic minorities; for instance, [Longley, Adnan, and Lansley \(2015\)](#) found that Twitter users in London are predominantly white British people and young males.

Population bias has also been identified in other cases, such as the US Twitter users ([Malik, Lamba, Nakos, & Pfeffer, 2015](#)) and Chinese Weibo users ([Yuan, Wei, & Lu, 2018](#)). The implication is that public opinion obtained from social media data, although in near real-time, are likely to be over-represented by certain groups in London (e.g., the young and males). The topics and attitudes regarding the reopening policy might be biased in relation to the Tweeting population in London. However, considering the cost- and time-efficiency advantages of social media data and the near real-time response from the aggregate dataset, this type of data source still retains substantial potential in sensing public opinion and in supporting policy making. To transcend the above-mentioned limitations of social media data, future studies can identify the socio-demographic (e.g., gender, age and ethnicity) make-up of social media users prior to data mining processes involving sentiment analysis and topic extraction. By revealing the composition of the data contributors, the results obtained from social media data can be more targeted, providing a more robust analysis for specific groups, thereby mitigating over-representations of populations.

Regarding limitations in the method, the techniques for sentiment analysis and topic modelling have potential to be upgraded in further studies, utilising more advanced algorithms. Both AFINN analysis and the LDA model are widely used and provide relatively accurate results. However, more nuanced emotions (such as anger, depression and joy) underlying the textual content of each tweet cannot be comprehensively identified through the AFINN method. To conduct more comprehensive sentiment analysis, further studies are recommended to add emotion detection capabilities. By further analysing the main topics and terms of tweets that express a specific emotion (e.g. anger), we can understand the underlying reasons that precipitate or catalyse a particular public attitude, which can further assist in developing a more dynamic and responsive mode of urban governance. In addition, the LDA model may

be computationally resource-intensive when processing a large amount of corpus ([Bhat, Kundroo, Tarray, & Agarwal, 2020](#)). Future studies can make use of more efficient topic modelling, such as deep learning enhanced topic modelling techniques, and thereby reduce the processing time.

6. Conclusion

This study focuses on public opinion and sentiment during the road to recovery from COVID-19 lockdown restrictions by analysing COVID-19 Twitter discussions regarding the reopening measures in Greater London utilising text mining approaches. Our findings reveal that public attention towards COVID-19 related discussions was high, but that public attention was decreasing as time went by, especially within the last four months. Among the COVID-19 discussions, tweets that were relevant to reopening measures were in a small proportion (less than 5 %). We find that reopening-related discussions reached a peak when the government announced its provisional reopening timeline and plans, which is to say that interest peaked prior to implementations. The sentiment results indicate that COVID-19 related posts were negative and consistently lower than general sentiment across the city. In terms of reopening debates, the sentiment scores were also negative, but relatively more positive than COVID-19 related discussion.

The topic modelling results further provide detailed insights into public opinion towards different reopening measures. We observed that people are more sensitive to the earlier reopening measure. For the reopening of schools, as identified in the previous section, Londoners worried about a rushed reopening implemented by government, with discussion highlighting the warnings from SAGE and the shadow cabinet. Regarding the non-essential shops reopening, the main discussions covered the topics related to being welcome to visit these shops, the confidence of the retail industry, as well as concerns about the practices of social distancing restrictions. When it comes to the reopening of hospitality, the social-economic recovery brought by this measure was discussed intensively. These results are cross-validated by the official opinion survey regarding COVID-19, confirming the robustness of this study.

We see the potential of our findings and methodology. First, our study can contribute to the existing scientific understanding of COVID-19 and post-crisis recovery. Advanced text mining methods are used to mine social media data and thereby facilitate understanding of government responses and public attitudes towards the pandemic. Second, the key findings herein have the potential to assist the recovery process in Greater London by enhancing public engagement and listening to the voice of the public. Text-mining results reveal the main focuses, concerns, and preferences prevalent during the reopening processes. Regional and national recovery strategies and timelines can be adapted to the main attitudes and concerns of the public. Third, big data analytics and the potential of social media data are discussed both in the literature review and throughout our empirical study. Social media data provides a larger data sample and diverse perspectives in terms of reopening measures during the pandemic crisis and in the aftermath. The dataset and advanced analytical tools (such as sentiment analysis and topic modelling) allow researchers, policymakers and governors to gain valuable, real-time insights. A more adaptive, flexible, and

inclusive urban governance could then be built for a wide range of cities, tailored to their road to recovery.

CRedit authorship contribution statement

Yiqiao Chen: Methodology, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Haifeng Niu:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Elisabete A. Silva:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study. *Journal of Medical Internet Research*, 22.
- Ahmed, M. E., Rabin, M. R. I., & Chowdhury, F. N. (2020). COVID-19: Social media sentiment analysis on reopening. arXiv preprint arXiv:2006.00804.
- Alexander, D. E. (2014). Social media in disaster risk reduction and crisis management. *Science and Engineering Ethics*, 20, 717–733.
- Al-Shabi, M. A. (2020). Evaluating the performance of the most important lexicons used to sentiment analysis and opinions mining. *International Journal of Computer Science and Network Security*, 20, 51–57.
- Bhat, M. R., Kundroo, M. A., Tarray, T. A., & Agarwal, B. (2020). Deep LDA : A new way to topic model. *Journal of Information & Optimization Sciences*, 41, 823–834.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Blumer, H. (1948). Public opinion and public opinion polling. *American Sociological Review*, 13, 542–549.
- Brandenburg, J. E., Holman, L. K., Apkon, S. D., Houtrow, A. J., Rinaldi, R., & Sholas, M. G. (2020). School reopening during COVID-19 pandemic: Considering students with disabilities. *Journal of Pediatric Rehabilitation Medicine*, 13, 425–431.
- Cabinet Office. (2020). Our plan to rebuild: The UK Government's COVID-19 recovery strategy. Retrieved from <https://www.gov.uk/government/publications/our-plan-to-rebuild-the-uk-governments-covid-19-recovery-strategy>. (Accessed 30 July 2020).
- Capela, F. D., & Ramirez-Marquez, J. E. (2019). Detecting urban identity perception via newspaper topic modeling. *Cities*, 93, 72–83.
- Chen, Y., Silva, E. A., & Reis, J. P. (2021). Measuring policy debate in a regrowing city by sentiment analysis using online media data: A case study of Leipzig 2030. *Regional Science Policy & Practice*, 13(3), 675–692.
- Feezell, J. T. (2018). Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era. *Political Research Quarterly*, 71, 482–494.
- Gibbons, J., Nara, A., & Appleyard, B. (2018). Exploring the imprint of social media networks on neighborhood community through the lens of gentrification. *Environment and Planning B-Urban Analytics and City Science*, 45, 470–488.
- GLA. (2020). About London recovery structures. Retrieved from <https://www.london.gov.uk/coronavirus/londons-recovery-coronavirus-crisis/london-recovery-board>. (Accessed 30 July 2020).
- GLA. (2020). London's recovery from the coronavirus crisis. Retrieved from <https://www.london.gov.uk/coronavirus/londons-recovery-coronavirus-crisis>. (Accessed 30 July 2020).
- GLA Opinion Research. (2020). COVID-19 online diary - weeks 1 and 2. Retrieved from London Datastore <https://data.london.gov.uk/dataset/COVID-19-online-diary>. (Accessed 15 April 2021).
- GLA Opinion Research. (2020). COVID-19 online diary: Summary report. Retrieved from London Datastore <https://data.london.gov.uk/dataset/COVID-19-online-diary>. (Accessed 15 April 2021).
- Górska, A., Dobija, D., Grossi, G., & Stanisewska, Z. (2021). Getting through COVID-19 together: Understanding local governments' social media communication. *Cities*, 121, Article 103453.
- Guo, Y., Chen, J., & Liu, Z. (2022). Government responsiveness and public acceptance of big-data technology in urban governance: Evidence from China during the COVID-19 pandemic. *Cities*, 122, Article 103536.
- Hochtl, J., Parycek, P., & Schollhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, 26, 147–169.
- Hollander, J. B., & Renski, H. (2017). Measuring urban attitudes embedded in microblogging data: Shrinking versus growing cities. *Town Planning Review*, 88, 465–490.
- Isoaho, K., Gritsenko, D., & Makela, E. (2021). Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal*, 49, 300–324.
- Jasiński, A. (2021). COVID-19 pandemic is challenging some dogmas of modern urbanism. *Cities*, 121, Article 103498.
- Jiang, H. C., Qiang, M. S., & Lin, P. (2016). Assessment of online public opinions on large infrastructure projects: A case study of the three gorges project in China. *Environmental Impact Assessment Review*, 61, 38–51.
- Kaufman, B. G., Whitaker, R., Lederer, N. M., Lewis, V. A., & McClellan, M. B. (2020). Comparing associations of state reopening strategies with COVID-19 burden. *Journal of General Internal Medicine*, 35, 3627–3634.
- Kaushal, J., & Mahajan, P. (2021). Asia's largest urban slum-dharavi: A global model for management of COVID-19. *Cities*, 111, Article 103097.
- Kearney, M. W. (2020). Rtweet: Collecting and analyzing twitter data. *Journal of Open Source Software*, 4(42), 1829.
- Lin, R., Lin, S., Yan, N., & Huang, J. (2021). Do prevention and control measures work? Evidence from the outbreak of COVID-19 in China. *Cities*, 118, Article 103347.
- Liu, Z. L., Lin, S. N., Shen, Y., & Lu, T. T. (2021). Collaborative neighborhood governance and its effectiveness in community mitigation to COVID-19 pandemic: From the perspective of community workers in six chinese cities. *Cities*, 116, Article 103274.
- Longley, P. A., Adnan, M., & Lansley, G. (2015). The geotemporal demographics of twitter usage. *Environment and Planning A*, 47, 465–484.
- Loper, E., & Bird, S. (2002). NLTK: the natural language toolkit. Retrieved from. arXiv preprint cs/0205028.
- Malik, M., Lamba, H., Nakos, C., & Pfeffer, J. (2015). Population bias in geotagged tweets. In , Vol. 9. *Proceedings of the international AAAI conference on web and social media*.
- Martin, A., Julian, A. B. A., & Cos-Gayon, F. (2019). Analysis of twitter messages using big data tools to evaluate and locate the activity in the city of Valencia (Spain). *Cities*, 86, 37–50.
- Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. (2013). Is the sample good enough? Comparing data from twitter's streaming api with twitter's firehose. In , Vol. 7. *Proceedings of the international AAAI conference on web and social media*.
- Nielsen, F. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.
- Niu, H., & Silva, E. A. (2020). Crowdsourced data mining for urban activity: Review of data sources, applications, and methods. *Journal of Urban Planning and Development*, 146(2), Article 04020007.
- O'Dowd, A. (2020). COVID-19: Papers justifying government's plans to reopen schools are "inconclusive," say union bosses. British Medical Journal Publishing Group.
- ONS. (2020). Coronavirus and the social impacts on the countries and regions of Britain: April 2020. Retrieved from. Office for National Statistics <https://www.ons.gov.uk/releases/coronavirusandthesocialimpactsinbritishregionsapril2020>. (Accessed 1 July 2020).
- Pereira, G. V., Parycek, P., Falco, E., & Kleinhans, R. (2018). Smart governance in the context of smart cities: A literature review. *Information Policy*, 23, 143–162.
- Preskorn, S. H. (2020). The 5% of the population at high risk for severe COVID-19 infection is identifiable and needs to be taken into account when reopening the economy. *Journal of Psychiatric Practice*, 26, 219–227.
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks*. Citeseer.
- Rufai, S. R., & Bunce, C. (2020). World leaders' usage of twitter in response to the COVID-19 pandemic: A content analysis. *Journal of Public Health*, 42, 510–516.
- Samuel, J., Rahman, M., Ali, G., Samuel, Y., & Pelaez, A. (2020). Feeling like it is time to reopen now? COVID-19 new normal scenarios based on reopening sentiment analytics. arXiv preprint arXiv:2005.10961.
- Sheikh, A., Sheikh, A., Sheikh, Z., & Dhami, S. (2020). In , *Journal of Global Health* Reopening schools after the COVID-19 lockdown (p. 10).
- Shen, C. H., Chen, A. F., Luo, C., Zhang, J. W., Feng, B., & Liao, W. (2020). Using reports of symptoms and diagnoses on social media to predict COVID-19 case counts in mainland China: Observational infoveillance study. *Journal of Medical Internet Research*, 22.
- Sievert, C., & Shirley, K. (2014). LDavis: A method for visualizing and interpreting topics. Retrieved from. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces* (pp. 63–70).
- Silva, E. A., Liu, L., Kwon, H. R., Niu, H., & Chen, Y. (2020). Hard and soft data integration in geocomputation: Mixed methods for data collection and processing in urban planning. *Handbook of Planning Support Science*.
- Sutton, J., Renshaw, S. L., & Butts, C. T. (2020). COVID-19: Retransmission of official communications in an emerging pandemic. *Plos One*, 15.
- Taecharungroj, V., & Mathayomchan, B. (2020). The big picture of cities: Analysing flickr photos of 222 cities worldwide. *Cities*, 102.
- Tan, S. B., Chiu-Shee, C., & Duarte, F. (2022). From SARS to COVID-19: Digital infrastructures of surveillance and segregation in exceptional times. *Cities*, 120, Article 103486.
- Tsao, S.-F., Chen, H., Tisseverasinghe, T., Yang, Y., Li, L., & Butt, Z. A. (2021). What social media told about us in the time of COVID-19: A scoping review. Retrieved from. arXiv preprint arXiv:2101.01688.
- UN-Habitat. (2020). World cities report 2020: The value of sustainable urbanization. Available online: Nairobi, Kenya: UN-Habitat https://unhabitat.org/sites/default/files/2020/10/wcr_2020_report.pdf.
- Vayansky, I., Kumar, S. A. P., & Li, Z. L. (2019). An evaluation of geotagged twitter data during hurricane Irma using sentiment analysis and topic modeling for disaster resilience. In *2019 IEEE international symposium on technology and society (Istas)*.
- Vest, J. R., Blackburn, J., & Yeager, V. A. (2021). Challenges in translating national and state reopening plans into local reopening policies during the COVID-19 pandemic. *Public Health Reports*, 136, 127–131.

- Wlezien, C. (2017). Public opinion and policy representation: On conceptualization, measurement, and interpretation. *Policy Studies Journal*, 45, 561–582.
- Wu, G., Deng, X., & Liu, B. (2022). Managing urban citizens' panic levels and preventive behaviours during COVID-19 with pandemic information released by social media. *Cities*, 120, Article 103490.
- Xu, Z., Liu, Y. H., Yen, N. Y., Mei, L., Luo, X. F., Wei, X., & Hu, C. P. (2020). Crowdsourcing based description of urban emergency events using social media big data. *IEEE Transactions on Cloud Computing*, 8, 387–397.
- Yao, Z. R., Yang, J. Y., Liu, J. L., Keith, M., & Guan, C. H. (2021). Comparing tweet sentiments in megacities using machine learning techniques: In the midst of COVID-19. *Cities*, 116.
- Yuan, Y. H., Wei, G. X., & Lu, Y. M. (2018). Evaluating gender representativeness of location-based social media: A case study of weibo. *Annals of GIS*, 24, 163–176.
- Zhang, X. Q., Ji, Z., Zheng, Y. Q., Ye, X. Y., & Li, D. (2020). Evaluating the effect of city lock-down on controlling COVID-19 propagation through deep learning and network science models. *Cities*, 107.
- Zhou, X. J., Tao, X. H., Rahman, M. M., & Zhang, J. (2017). Coupling topic modelling in opinion mining for social media analysis. In *2017 IEEE/WIC/Acm international conference on web intelligence (Wi 2017)* (pp. 533–540).