Biases in Using Social Media Data for Public Health Surveillance: A Scoping Review

Yunpeng Zhao[1a], Xing He[1a], Zheng Feng[1], Sarah Bost[1], Mattia Prosperi[2], Yonghui Wu[1], Yi Guo[1], Jiang Bian[1b]

[1]Health Outcomes & Biomedical Informatics, College of Medicine, University of Florida, Gainesville City, FL, United States

[2]Department of Epidemiology, University of Florida, Gainesville City, FL, United States

Email: Yunpeng Zhao – yup111@ufl.edu; Xing He – hexing@ufl.edu; Zheng Feng - fengzheng@ufl.edu; Yi Guo – yiguo@ufl.edu; Mattia Prosperi – mattia@ufl.edu; Yonghui Wu – yonghuiwu@ufl.edu; Jiang Bian – bianjiang@ufl.edu

[a]Yunpeng Zhao, MS and Xing He, MS contributed equally, co-first authors

[b]**Corresponding author:** Jiang Bian, PhD; bianjiang@ufl.edu.

Affiliation: Health Outcomes & Biomedical Informatics, University of Florida

Address: 2197 Mowry Road, 122 PO Box 100177 Gainesville, FL 32610-0177

Phone Number: (813)573-3122

Keywords: social media, bias, public health surveillance

ABSTRACT

Objectives: A landscape scan of the methods that are used to either assess or mitigate biases when using social media data for public health surveillance, through a scoping review.

Materials and Methods: Following best practices, we searched two literature databases (i.e., PubMed and Web of Science) and covered literature published up to July 2021. Through two rounds of screening (i.e., title/abstract screening, and then full-text screening), we extracted study objectives, analysis methods, and the methods used to assess or address the different biases from the eligible articles.

Results: We identified a total of 2,856 articles from the two databases. After the screening processes, we extracted and synthesized 20 studies that either assessed or mitigated biases when leveraging social media data for public health surveillance. Researchers have tried to assess or address several different types of biases such as demographic bias, keyword bias, and platform bias. In particular, we found 11 studies that tried to measure the reliability of the research findings from social media data by comparing them with other data sources.

Discussion and Conclusion: We synthesized the types of biases and the methods used to assess or address the biases in studies that use social media data for public health surveillance. We found very few studies, despite the large number of publications using social media data, considered the various bias issues that are present from data collection to analysis methods. Overlooking bias can distort the study results and lead to unintended consequences, especially

in the field of public health surveillance. These research gaps warrant further investigations more systematically. Strategies from other fields for addressing biases can be introduced for future public health surveillance systems that use social media data.

BACKGROUND AND SIGNIFICANCE

Social media platforms are internet places for people to connect. Social media users often voluntarily discuss and share their health-related experiences, such as their concerns about contracting certain diseases or vaccinations.[1,2] These health-related posts on various social media platforms bring new opportunities for public health surveillance. There are different focuses of using social media data for public health surveillance, such as (1) disease surveillance,[3,4] (2) pharmacovigilance,[5,6] (3) misinformation surveillance,[7] (4) surveillance of human mobility and health behavior of a population, some of which use location-based social networks.[8–10] Nevertheless, the nature of social media data and associated analysis methods are very different from those that are used in traditional public health surveillance systems. Traditionally, surveillance systems can be classified into either active or passive surveillance based on the way they collect the data. For active surveillance systems, data are collected through active outreach such as from surveys that ask questions of specific public health-related events, where different sampling or weighting strategies are often used to create results that can well represent the target population.[11] For passive surveillance systems, data are passively collected such as relying on reports by health care providers.[12] Social media data are often used in passive surveillance systems, where they passively monitor organic social media posts to identify events of interest.[3] Nevertheless, it is critical to recognize the unique challenges of dealing with the various potential biases in using social media data for public health surveillance. A well-known example of harmful consequences when biases are ignored is Google Flu Trends' failure of making accurate predictions using internet search data.[13] Even though Google Flu Trends did not use social media data, many of the potential biases are commonly inherent in surveillance using internet data, such as representativeness, confounding of search terms, and lack of case validation.[14] On a high level, we can generally categorize the biases from their sources: rising (1) from the data itself, and/or (2) from the methods used when processing and analyzing the data.

Biases inherent in the social media data

"*Data bias*" is the biases that comes from the inherent properties of social media data. For example, social media data may not be representative of the general population of interest, while representativeness is often a key desired feature of an ideal surveillance system. Firstly, the demographics of social media users are not only different from the real-world populations but also different across social media platforms. An early study from the Pew Research Center discovered that TikTok and Instagram have more female users than male users, while male users are more prominent on Twitter.[15] Certain populations (e.g., younger adults and those that are more comfortable with technology) are more prevalent on social media platforms in part due to the characteristics of the specific subpopulations but also the particular design and marketing strategies of the different social media platforms.[15] Compounding this issue is that social media platforms either do not collect user demographics explicitly such as Twitter or do not make them available, for the right reason of protecting user privacy, such as Facebook, which makes it difficult to use traditional methods (e.g., raking [16]) to generalize the findings from social media data to the general populations. Some researchers have attempted to infer user demographics from other contextual features that are available about the social media user to address some of these issues. For instance, Culotta et al. (2015) created a machine learning classifier to identify Twitter users' ethnicity, gender, and political preference based on whom they follow.[17] However, some of the demographic attributes (e.g., age) are still difficult to extract. Nguyen et al. (2014) found that older Twitter users are often predicted to be younger using features derived from their Twitter posts, introducing additional biases if used to adjust for represenativeness.[18] Lastly, social media data may contain information posted by fake user accounts or bots. A recent study found that bots contributed to nearly half of the discussions about "*reopening America*" during the COVID-19 pandemic on Twitter.[19] For a surveillance system, it is important to identify and remove posts from bots or fake accounts.

Biases raised from the methods used in dealing with social media data

"*Method bias*" refers to the biases that come from the methods and procedures applied for the collection, processing, and analysis of the social media data. For example, most social media studies identify and collect sample datasets by using keywords and hashtags, depending on the interfaces provided by individual social media platforms. Such keyword-based searches may lead to biased samples (e.g., not representative of the topic of interest) and introduce noises (i.e., data irrelevant to the topic of interest) due to the ambiguity of the keywords. Using keywords may also have a low recall, since it is difficult to identify all the relevant keywords and the vocabulary used on social media are often different from those used in formal writing and evolves rapidly (e.g., new slang terms continuously being invented). Thus, the choice of keywords (and hashtags, in the case of Twitter) determines both the precision and recall of the retrieved dataset in terms of its relevance to the topic of interest. Existing studies have shown that poorly designed search queries can introduce more biases.[20] Secondly, regardless of the methods used for data collection, the sample data retrieved from social media platforms is only a fraction of all relevant data. Social media platforms such as Twitter provide application program interfaces (APIs) for data accessing purposes, but with restrictions on query length, data volume, and data request frequency.[21] Lastly, different from other traditional passive surveillance data sources such as structured, coded data from electronic health records (EHRs), social media data are often unstructured free-text data, where natural language processing (NLP) methods are frequently used (e.g., text classifiers, sentiment analysis, and topic modeling[22–24]). These NLP methods can introduce biases (e.g., misclassification errors introduced by the classifiers). Further, data preprocessing procedures, often a necessary step in the NLP pipeline, can also introduce biases. Standard text normalization methods, such as spelling corrections, lemmatization, and stemming, can potentially alter the meaning of original words or phrases. For example, stemming the words *"flying"* and *"flies"* (i.e., the insects) will lead to an identical representation, i.e., *"fly."* These data preprocessing methods may also lead to radically different results of the downstream NLP models. For example, topic modeling techniques can yield different results depending on the choices made in the different pre-processing steps for textual data.[25]

There are growing concerns of both the data and method biases when using social media data for public health surveillance. Overlooking biases can distort the study results and lead to unintended consequences. Even though the awareness is high,[3] there is limited work on strategies to either assess (e.g., quantify) or mitigate the biases. Thus, our goal of this study is to conduct a landscape scan of the methods that are used to either assess or mitigate biases when using social media data for public health surveillance, through a scoping review the literature. To do that, we aim to answer the following two research questions (RQ):

- RQ1: What are the existing data analysis methods (e.g., machine learning models for classification) used in social media studies related to public health surveillance?
- RQ2: What are the existing methods used to assess and/or address bias in social media studies related to public health surveillance?

Through answering these two RQs, we will identify research gaps from social media studies in the field of public health surveillance. To the best of our knowledge, there are no existing reviews focusing on this topic, i.e., biases in social media studies for public health surveillance. Similar discussions in review literature can only be found on biases of general social media or social network studies[26] or on biases of public health surveillance using traditional data sources such as electronic health records.[27]

MATERIALS AND METHODS

Literature search strategies

This scoping review follows the best practices and uses the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Through a systematic search of two representative literature databases (i.e., PubMed and the Web of Science), we identified relevant articles that assessed and/or addressed data and method biases in using social media data for public health surveillance published by July 6th, 2021. Supplement **Appendix A** shows the search strategies we used, which contains three groups of keywords: (1) public health surveillance-related, (2) social media related; and (3) bias-related. The initial social media and bias-related keywords were built upon a survey paper that discusses biases in general social media studies;[28] and we developed the public health surveillance-related keywords through a manual screening of relevant MeSH terms and samples of relevant studies. Through this process, we found some social media studies often use machine-/deep learning (ML/DL) methods to filter out irrelevant information or bot accounts from social media data, which is a way of reducing the biases introduced by these nosies. These studies are less likely to mention terms related to "*bias*" but can also be highly relevant to our RQs; we thus also included ML and bot-related keywords to the bias-related keyword group.

Eligibility criteria

We drafted the initial inclusion and exclusion criteria through group discussions and conducted two rounds of initial exercises of title and abstract screening to train the reviewers and refine the eligibility criteria. The final inclusion criteria are: (1) studies that use data (i.e., any data types, including text, images, and videos) generated from social media platforms, (2) studies that are related to public health surveillance, and (3) the studies should have evaluated and/or addressed/mitigated the biases in the social media data itself (i.e., data bias) and/or the analysis methods (i.e., method bias) that used to process the social media data. We excluded studies that: (1) are not written in English, (2) are review, opinion, and perspective papers, and (3) not related to analysis of social media data for public health surveillance (e.g., use the social media platforms for recruitment).

Article screening process

Following the PRISMA guideline, we first removed duplicate records across the two literature databases and conduced title and abstract screening based on our inclusion and exclusion criteria. During this process, we also iteratively refined the eligibility criteria. For the articles that passed title/abstract screening, we conducted a full-text screening. In both title/abstract and full-text screenings, two reviewers (YZ and XH) performed the screening independently, and conflicts were resolved by a third reviewer (JB).

Data extraction from the articles

We developed a data extraction form iteratively during the full-text screening phase with a focus on information related to the objective of each study and how data and method biases were assessed and/or addressed. For each study, we extracted: (1) the outcomes of interest (e.g., conditions, diseases, or adverse events) , (2) the social media data sources (e.g., Twitter), (3) the data analysis methods (e.g., ML-based classifier), and (4) whether the study addressed data and/or method biases and if so the types of the bias that were addressed.

RESULT

A total of 2,856 articles were identified from the two literature databases. After removing duplicates, 2,193 articles were left for title and abstract screening; from which, 2,159 articles were deemed ineligible because they either do not use social media data for surveillance or do not explicitly assess or address data or method biases according to our eligibility criteria. For articles that its eligibility is unclear from the title and abstract alone, we conservatively kept the article for full text screening. We further screened the full text of the remaining 34 articles and removed 12 articles that have no bias evaluation and 2 articles that are not related to public health surveillance. Finally, 20 articles remained eligible for data extraction. **Figure 1** shows the PRISMA flow diagram of our review process.

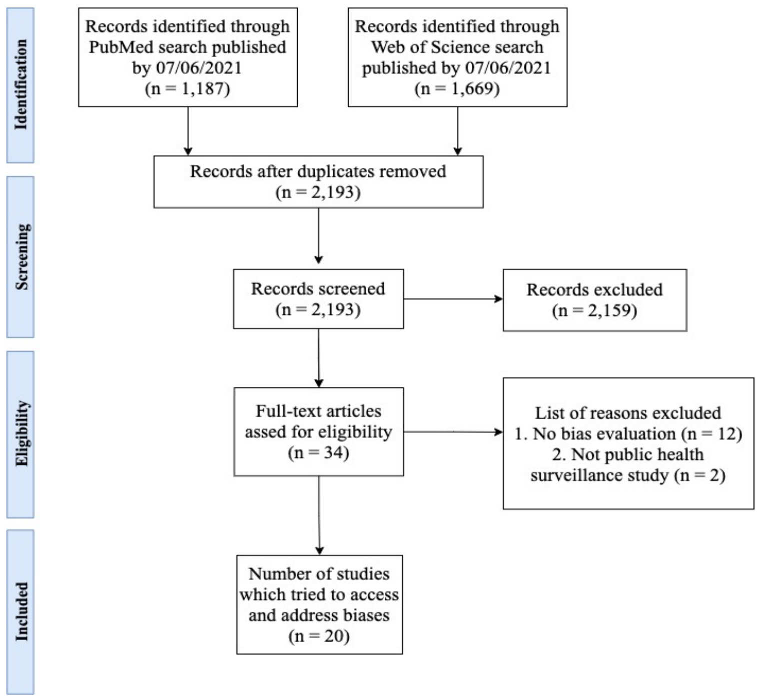

Figure 1. PRIMSA flow diagram of the literature review process.

Overview of the included studies

Among the 20 articles included for data extraction, 7 different social media platforms were used: Twitter (n = 15), Facebook (n = 2), Yelp (n = 1), Weibo (n = 1), YouTube (n = 1), Instagram (n = 1), and web forums (n = 1). Twitter is the most popular data source for public health surveillance studies using social media data. There are 3 studies that used data from multiple social media platforms: (1) Audeh et al. (2020) used data from 21 French web forums to detect drug mentions;[29] (2) Elkin et al. (2020) manually evaluated vaccination-related contents from YouTube and Facebook;[30] and (3) Jaidka et al. (2020) estimated geographic well-being by using data from Twitter and Facebook.[31] Except for two studies[30,32] that manually coded the content of videos and images from YouTube and Instagram, respectively, all the other studies analyzed textual data from social media.

The outcomes of interest in the 20 articles are (1) disease surveillance (n= 14; e.g., infectious disease), (2) pharmacovigilance (n=7; i.e., adverse events, drug use/misuse, and vaccination), (3) public's attitudes or behaviors (n=4), and (4) others (n=2; e.g., general well-being). **Table 1** shows the number of studies by the outcome of interest. Disease surveillance (n = 14) is the most prevalent use case for public health surveillance using social media data, including infectious diseases (n = 10), chronic diseases (n = 2), and mental health (n = 3). Two of the 10 infectious disease studies focused on the current pandemic of Coronavirus disease 2019 (COVID-19). Note that some studies studied multiple diseases or multiple outcomes. For example, Yang et al (2016)[33] created a general-purpose platform and discussed three different use cases: influenza outbreaks (i.e., infectious disease), public responses to Ebola outbreak (i.e., attitudes and opinions), and online discussion of (medical) marijuana (i.e., drug use).

Table 1. Summary of the outcomes of interest among the 20 included studies.

Outcomes	**Specific outcomes**	**Number of studies**	**Reference**
Disease surveillance	Infectious diseases (e.g., COVID)	10	33–42,
	Chronic diseases	2	42,43
	Mental health (e.g., depression)	3	32,42,44
Pharmacovigilance	Adverse Event	2	29,45

	Vaccine	1	30
	Drug use/misuse (e.g., opioid)	4	29,33,46,47
Public's attitudes or behavior	Attitudes and behavior (e.g., opinions, alcohol consumption)	4	33,36,37,39
Other	General well-being	2	31,48

RQ1: What are the existing data analysis methods used in social media studies related to public health surveillance?

To answer RQ1, we extracted the analysis methods used in the 20 studies and categorized these into 3 groups: (1) classification models, including both ML-based classification (e.g., Aslam et al. (2014) implemented a support vector machine to identify laypeople' flu-related tweets [34]) and rule-based classification (e.g., Yang et al. (2016) adopted simple rules that remove retweets and tweets with URLs to remove irrelevant information[33]). Note that we considered ML- or dictionary-based sentiment analysis into this category as well; (2) content analysis that includes both algorithmic text clustering or topic modeling methods (e.g., Massey et al. (2021) explored discussions topics from Twitter data on the topic of COVID-19 using topic modeling[37]) and manual content analysis (e.g., McCosker et al. (2020) developed a manual coding approach to explore depression-related contents on Instagram[32]); and (3) correlation analysis that includes simple correlation measures (e.g., Jayawardhana et al. (2019) validated the influenza rate estimates from social media data with hospitalization records issued by Ohio Departmetn of Health[35]) and regression analysis (e.g., Alessa et al. (2019) used linear regression with flu-related tweets to estimate flu-rate[40]). **Table 2** shows the number of studies by analysis method. Note that some studies employed multiple methods.

Table 2. The number of studies by analysis method.

Categories	Methods	Number of studies*	Reference
Classification models	Machine learning-based classification/sentiment analysis	12	[30,33–36,38–40,43–45,48]
	Rule-based classification/dictionary-based sentiment analysis	4	[31,33,39,46]
Content analysis	Manual content analysis	2	[29,32]
	Text clustering/topic modeling	3	[37,41,47]
Correlation analysis	Simple correlation measures	10	[31,32,34–37,40,41,46,48]
	Regression analysis	2	[39,40]
*Note that some studies used multiple analysis methods.			

RQ2: What are the existing methods used to assess and/or address bias in social media studies related to public health surveillance?

To answer RQ2, we first summarized the types of biases that were discussed in the 20 studies based on existing literature on the topic of bias in public health surveillance.[49,50] Nevertheless, there is no standard classification of biases and the definition of each bias; and it is often difficult to draw clear boundaries between different bias terms and their normative connotations. Table 3 shows the summarization along with the definition or example of the specific bias type, the methods used for assessing or mitigating the bias, and associated studies. Out of the 20 studies, 10 of them (i.e., some studies addressed multiple biases) discussed three types of biases: (1) demographic bias (n=3), (2) keyword bias (n=8), and (3) platform bias (n=1), which all related to selection bias. Most studies focused on discussing the biases of the social media data, while a few (i.e., 8 articles that discussed keyword bias) addressed the biases introduced by the methods used to collect, process, or analyze the data. Even the 8 articles that are related to keyword bias have focused their discussions on how issues concerning the choice and use of certain keywords would affect the sample data (i.e., data bias due to data collection or processing methods used). There is no study that discussed how analysis methods would introduce biases in the study results explicitly.

Table 3. Summary of the bias types and methods to assess or mitigating the bias in the 20 articles.

Type of the bias in public health surveillance literature	**Example/definition**	**# of studies**	**Methods for assessing**	**Methods for mitigating**	**Studies**	**Data bias or Method Bias**
Demographic bias	The demographics of the social media user shifts from the general population	3	NA	Stratifying social media users based on the demographic distributions	[31,32,48]	**Data**
Keyword bias	The use of keywords to extract sample data may introduce noises as the keywords may be ambiguous (e.g., misspelling or slang words)	8	Manual analysis	(1) Machine learning-based filtering (2) Rule-based filtering	[31–34,39,43–45]	Data/Method
Platform bias	Differences across platforms due to platform characteristics (e.g., the ranking algorithm it used)	1	Manual analysis	NA	[30]	Data
Unclassified*		10	Regression or correlation analysis	NA	[29,35–37,40–43,46,47]	Data
* Studies that cannot be mapped to existing types of biases from public health surveillance literature; however, some of the studies in this category compared their social media results with other data sources, thus, in a way assessed the biases of the study results. See Table 4 for details on those individual studies.						

From the 20 studies, we found 3 discussed demographic bias. Iacus et al. (2020)[48] and Jaidka et al. (2020)[31] attempted to mitigate demographic bias by stratifying Twitter users based on their geographic distributions to get representative measurements of users' general well-being from Twitter data, while Weeg et al. (2015) [32] found that the correlation between findings from social media data and the results from a national survey was significantly increased after stratifying Twitter users by demographics. Eight out of 20 studies have targeted keyword bias. For example, Mowery et al. (2017) assessed how accurately the depression-related keywords could identify depression-related tweets by manually reviewing a sample of tweets for each keyword;[44] and Culotta et al. (2013) tested both rule-based (i.e., keyword-based) approach and machine learning-based approach to identify relevant tweets and used the volume of the identified tweets to estimate flu rates and alcohol sales volume from Twitter data.[39] We found only 1 article that attempted to assess social media platform bias. Elkin et al. (2020)[30] manually evaluated vaccine-related content from YouTube and Facebook; and they found more negative vaccine-related content on Facebook than YouTube.

However, the rest 10 out of the 20 studies addressed the overall data or method bias question but cannot be classified into the 3 types of biases described above. Most of these studies (7 out of the 10) discussed the reliability of social media study results when potential bias exists by validating the results generated from social media data with external data sources. In fact, there is a total of 11 studies (4 from those that can be classified into the 3 types of biases described above) that compared social media results with external data sources, and we further list the specific validation methods and corresponding external data sources used in the 11 studies in

Table 4. We found data sources such as hospitalization records,[35] reports from the Centers for Disease Control and Prevention (CDC)[51,52] and surveys[32,48] are often used as the external validation datasets; and 9 out of 11 articles used simple correlation metrics to compare the results from social media data with the external data sources. At last, 3 studies[29,36,43] that cannot be classified, as they are general descriptive studies (e.g., Audeh et al. (2020)[29] identified the most frequently mentioned drugs in web forums and discussed the potential biases related to forum selection and the corresponding population representativeness).

Table 4. Social media public health surveillance studies that compared their results with external data sources.

Validation method	**Articles**	**Topic**	**External data source**
Simple correlation	Aslam et al. (2014)[34]	Seasonal influenza surveillance from Twitter	The morbidity and mortality weekly report by the CDC[52]
	Weeg et al. (2015)[32]	Disease mentions vs. prevalence from Twitter	Survey data by the Experian Marketing Services[53]
	Chary et al. (2017)[47]	Misuse of opioids estimation from Twitter	The national survey on drug usage and health[54]
	Jayawardhana et al. (2019)[35]	Influenza rate from Twitter	The hospitalization records by the Ohio Department of Health[55]
	Jaidka et al. (2020)[31]	Well-being distribution from Twitter	The Gallup-sharecare well-being index survey[56]
	Iacus et al. (2020)[48]	Well-being distribution from Twitter	Survey data from the Italian National Institute of Statistics (ISTAT)[57]
	Massey et al. (2021)[37]	COVID-19 case prediction using Twitter	The United States COVID-19 cases and

			deaths by the state over time reports by the CDC[51]
	Margus, et al. (2021)[41]	COVID-19 case prediction using Twitter	The COVID-19 dashboard by the Center for Systems Science and Engineering at Johns Hopkins University[58]
	Tacheva et al. (2021)[46]	Misuse of opioids estimation from Twitter	A wide range of online data for epidemiologic research by the CDC[59]
Regression analysis	Culotta et al. (2013)[39]	Influenza rates from Twitter	The reports from the US outpatient influenza-like illness surveillance network by the CDC[60]
	Alessa et al. (2019)[40]	Flu detection from Twitter	FluView by the CDC[61]

DISCUSSION

We summarized the existing studies that have discussed methods and strategies used to assess and/or mitigate data and method biases when using social media data for public health surveillance through a scoping review. Even though our initial literature database search identified a large number of records, only 20 articles eventually met our eligibility criteria that explicitly discussed either data or method biases when using social media for public health surveillance. Despite the great awareness of bias concerns, we found very few studies have explored this topic, and virtually no practical and systematic methods have been proposed to mitigate the various biases when using social media data. Although some studies have realized the potential biases, they failed to identify the specific types of biases and address them according to their properties. Only 10 studies further discussed biases in different types. Eleven out of the 20 studies discussed the reliability of study results when potential biases exist by comparing or validating the results with external, often more authoritative data sources such as those from the CDC. For studies that discussed and addressed biases of different types, there is a significant under-awareness of several types of biases and only a few types of the biases (Table 3) are unevenly discussed. Among the 20 studies we reviewed, 8 addressed keyword bias, 3 addressed demographic bias, and only 1 study addressed platform bias. Even though sample bias and misclassification errors are discussed extensively in existing literature on biases in public health surveillance studies[49,50] and in general social media studies,[28] we did not find any social media studies that addressed either sample bias or misclassification errors directly.

Based on our findings above and by exploring strategies of addressing biases that is used in studies on social media from fields other than public health surveillane,[28] we discuss 5 types of biases below and recommend more up-to-date tools for each type of the bias that can be considered for future public health surveillance system of using social media data as follows.

Demographic bias

Stratifying social media users based on their demographic distributions to get representative results from social media data is a useful approach. However, demographic information is unavailable on many social media platforms (e.g., Twitter), so that researchers often have to build models to infer those information.[62] Further, beyond simple demographics (e.g., age, gender, race, and ethnicity), researchers have been able to create models to infer other social media user attributes. For example, Daniel et al. (2015) tested support vector machine (SVM) and linear regression models to predict the income level of Twitter users.[63] Michael et al. (2011) used a SVM model to predict the political alignment of Twitter users based on their posts.[64] As many other kinds of sociodemographic information are possible to be extracted from social media data using advanced inference models, stratifying social media users by those attributes for public health surveillance can potentially provide more insights into the different subpopulations. Nevertheless, these inference models will also introduce misclassification errors because of the imperfection of these models.

Keyword bias

Both ML-based and rule-based methods are often applied to mitigate the keyword bias in studies we reviewed; nevertheless, Culotta et al. (2013) found that ML-based classifiers are more adept than rule-based methods for filtering out irrelevant information.[39] However, the irrelevant information introduced by ambiguous keywords is only one aspect of the keyword bias, where the coverage or completeness of all the potentially relevant data that the keywords can retrieve is another issue. When we developed search keywords for content filtering in our previous social media studies,[1,65] we considered keyword variations, misspellings, and vocabulary changes over time to collect as much relevant social media data as possible. Other approaches have been proposed outside of the topic of using social media data for public health surveillance. For example, Magdy et al. (2014) used an unsupervised machine learning approach to track dynamic topics and theme changes in Twitter data.[66] Nevertheless, without knowing the complete universe of the social media data space, the representativeness of the collected data and the generalizability of the study results are difficult to assess.

Platform bias

Different social media platforms often attract different user groups due to its unique characteristics, but the user behaviors on the different social media platforms might also be different. For example, Linkedin is designed to be business- and employment-oriented online service, where the posts and communications on Linkedin are mostly expressed more formally in a professional manner. On the other hand, other social media platforms such as Twitter and Facebook are geared toward making sharing content and communicating among families and friends, where the content posted are causal and less formal. The same user may behavior differently between social media platforms like Linkedin and Twitter. Only 1 article by Elkin et al. (2020)[30] discussed platform bias, where they found more negative vaccine-related content on Facebook than YouTube.

All the three types of biases discussed above (i.e., demographic bias, keyword bias, and platform bias) identified from the 20 studies are related to concept of selection bias in classific public health surveillance literature, where the bias is introduced by the selection of individuals (or their data) in a way that proper randomization is not achieved, which leads to an unrepresentative sample of the population intenteded to be studied. It is yet unclear how such selection bias can be addressed, given the inherent limitation of what data and information can be obtained from the different social media platforms. For example, Twitter although provides APIs for end-users to access public tweets, the sampling strategies that Twitter internally used for these API end-points are unknown to end-users, leading to difficult to migitating the introduced sampling bias. In studies outside of public health surveliance domain, a number of social media studies have discussed selection (and sampling) bias. For example, Morstatter et al. (2013) measured the representativeness of Twitter streaming API to the full archive dataset by comparing topics, geographic distributions, and networks of Twitter users between the two datasets.[67] Pfeffer et al. (2018) also used multi-crawlers to circumvent the API limits and discussed the possibility of collecting complete data using this strategy.[68]

Misclassification errors

Misclassification errors is a frequently discussed issue in public health surveillance literature and a common issue in studies that use classification models, while rule-based or ML-based algorithms are often used in social media studies to filter out irrelevant information. Nevertheless, none of the 20 studies discussed misclassification issues. Although most studies that use classification strategies have tested multiple classifiers and adept the one with the best performance, it is not sufficient. Classification models including ML-based classifiers are sensitive to biases that occurred in every step of the social media data processing and analysis pipeline (e.g., because of sample bias and sampling errors introduced in the data preprocessing step). For example, Thomas et al. (2020) pointed out that the representativeness of the training samples is extremely important to build reliable ML-based classification models.[69] To mitigate this bias, besides using more advanced ML models such as deep learners, we also need to solve and consider the biases that occurred in the data processing steps prior to building the actual ML models. Further, no model can achieve perfect performance; thus, systematic studies that provide insights on how biases in social media data would affect the performance of ML models and subsequently affect the final study results are warranted. For example, sensitivity analysis can be important to obtain confidence intervals when reporting the study results using models with varying performance.[70]

Final remarks

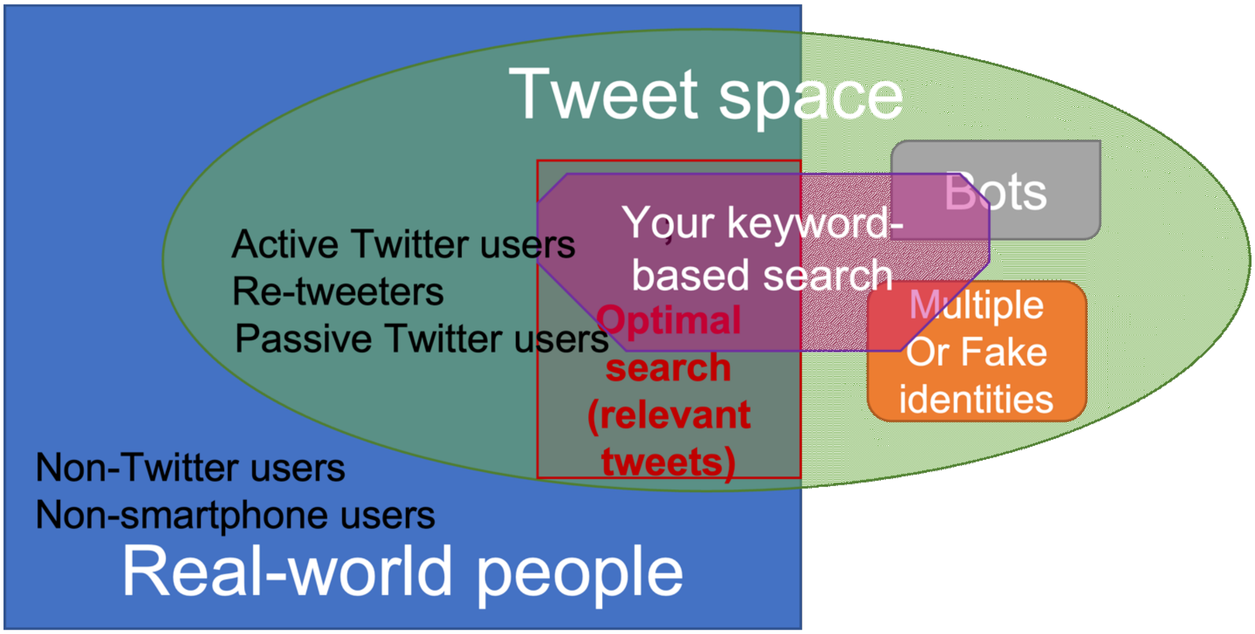

Figure 2. A conceptual view of the Tweet space.

For public health surveillance, biases not only exist inherently in the data itself but also can be introduced from the methods used to collect, process, and analyze the data. Nevertheless, issues around these biases are rooted in the question of whether the collected data samples represent the topic or individuals of interest. **Figure 2** shows a conceptual view of the social media data universe using Twitter as an example. The sample data that are collectable through Twitter APIs are not only just a subset of the optimal, desired search results, but also may contain irrelevant information from bots or fake accounts. Further, even within the relevant tweets, for the purpose of public health surveillance, we would need to consider the different user characteristics (e.g., active user vs. retweeters) that affect incidence and prevalence estimates that are critical in surveillance systems. Moreover, we have to keep in mind that even the entire Tweet space may not represent the real-world populations of interest, considering those disadvantaged populations who may not even have access to the internet or those are not Twitter users. So, an ultimate question that every researcher who aims to develop a public health surveillance system based on social media data should consider is whether the data source available can meet the surveillance question of interest? For example, social media data like tweets may be an excellent supplementary data source to identify novel symptoms of long COVID but may not be the right or sole data source for estimating the prevalence of COVID infections.

Ultimately, public health surveillance will need to be designed in ways that avoid or reduce the potential biases to "*guarantee*" the accuracy of the results and the robustness of the systems. It is critical to identify where these biases may come from, subsequently understand the issues that these biases bring to the studies of public health surveillance and address them with correct approaches considering the specific context of the studies. Based on all previous discussions, we suggest Table 5 for researchers to quickly evaluate their study goals and design their public health surveillance systems with the appropriate tools to address potential biases.

Table 5. Summary of the type of biases, the potential issues they will cause, and recommendations for addressing the biases.

Type of bias	Potential issues	Recommended approaches to address the corresponding bias
Demographic bias	Selection or sampling bias that will lead to an unrepresentative sample of the population intended to be studied	Stratify and social media users based on their demographic distributions.[58,59,60]
Keyword bias		(1) Use ML-based or rule-based filter to filter irrelevant information introduced by ambiguous keywords.[35] (2) Track dynamic topics and theme changes in Twitter data.[62]
Platform bias		Evaluate data property of different social media platforms and utilize the APIs provided by platforms.[63,64]
Misclassification errors	Affect the performance of models and subsequently affect the final study results	(1) Evaluate model representativeness when building ML/DL models.[65] (2) Use sensitivity analysis to obtain confidence intervals.[66]

LIMITATIONS

This review has two limitations. First, there is no standard taxonomy of bias and the definition of each bias term. The bias terms used in this review were summarized from the literature[49,50] on the topic of bias in public health surveillance. To provide the full picture of bias in using social media data, the taxonomy and clear definition of bias terms should be thoroughly developed from multiple fields in future research. Second, we only reviewed the articles on the topic of public health surveillance using social media data. Some biases are not unique to social media studies for public health surveillance but are shared among other social media data analysis studies. Further investigation on how bias in social medial data and analytic methods would affect study results should be systematically studied.

CONCLUSION

In this review, we identified the methods used to assess and address different biases in studies that use social media data for public health surveillance. We found that very few studies have been conducted on this topic, and we identified research gaps that warrant further investigations more systematically. The strategies of addressing bias in social media studies from other fields can be introduced for future public health surveillance systems that use social media data. But ultimately, researcher who aims to develop public health surveillance systems using social media should consider is whether the data source available can meet the surveillance question of interest.

FUNDING

This work was supported in part by NSF Award #1734134 and CDC Award U18DP006512.

CONTRIBUTORS

JB and YZ designed the initial concepts and framework for the proposed systematic review; YZ and HX extracted and summarized the bias taxonomy from public health surveillance literature; YZ and HX carried out the review and annotation process, and JB resolved annotation conflicts. YZ wrote the initial draft of the manuscript. JB, HX, FZ, YG, MP, YW provided critical feedback and edited the manuscript.

REFERENCES

1. Bian J, Zhao Y, Salloum RG, Guo Y, Wang M, Prosperi M, Zhang H, Du X, Ramirez-Diaz LJ, He Z, Sun Y. Using Social Media Data to Understand the Impact of Promotional

Information on Laypeople's Discussions: A Case Study of Lynch Syndrome. J Med Internet Res [Internet]. 2017 Dec 13;19(12):e414. Available from: http://dx.doi.org/10.2196/jmir.9266

2. Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and Demographic Correlates of Autism-Related Anti-Vaccine Beliefs on Twitter, 2009–15. Soc Sci Med [Internet]. 2017 Oct;191:168–175. Available from: http://dx.doi.org/10.1016/j.socscimed.2017.08.041 PMID: 28926775

3. Aiello AE, Renson A, Zivich PN. Social Media- and Internet-Based Disease Surveillance for Public Health. Annu Rev Public Health [Internet]. 2020 Apr 2;41:101–118. Available from: http://dx.doi.org/10.1146/annurev-publhealth-040119-094402 PMID: 31905322

4. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EHY, Olsen JM, Pavlin JA, Shigematsu M, Streichert LC, Suda KJ, Corley CD. Using Social Media for Actionable Disease Surveillance and Outbreak Management: A Systematic Literature Review. Braunstein LA, editor. PLoS ONE [Internet]. 2015 Oct 5;10(10):e0139701. Available from: http://dx.doi.org/10.1371/journal.pone.0139701

5. Sloane R, Osanlou O, Lewis D, Bollegala D, Maskell S, Pirmohamed M. Social media and pharmacovigilance: A review of the opportunities and challenges. British Journal of Clinical Pharmacology [Internet]. 2015;80(4):910–920. Available from: http://dx.doi.org/10.1111/bcp.12717

6. Pappa D, Stergioulas LK. Harnessing social media data for pharmacovigilance: a review of current state of the art, challenges and future directions. Int J Data Sci Anal [Internet]. 2019 Sep;8(2):113–135. Available from: http://dx.doi.org/10.1007/s41060-019-00175-3

7. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of Health Misinformation on Social Media: Systematic Review. Journal of Medical Internet Research [Internet]. 2021 Jan 20;23(1):e17187. Available from: http://dx.doi.org/10.2196/17187

8. Maher C, Ryan J, Kernot J, Podsiadly J, Keenihan S. Social media and applications to health behavior. Current Opinion in Psychology [Internet]. 2016 Jun 1;9:50–55. Available from: http://dx.doi.org/10.1016/j.copsyc.2015.10.021

9. Comito C, Forestiero A, Pizzuti C. Improving influenza forecasting with web-based social data. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) [Internet]. IEEE; 2018. Available from: http://dx.doi.org/10.1109/asonam.2018.8508563

10. Comito C. How COVID-19 information spread in US The Role of Twitter as Early Indicator of Epidemics. IEEE trans serv comput [Internet]. Institute of Electrical and Electronics Engineers (IEEE); 2021;1–1. Available from: http://dx.doi.org/10.1109/tsc.2021.3091281

11. Setia M. Methodology series module 5: Sampling strategies. Indian J Dermatol [Internet]. 2016;61(5):505. Available from: http://dx.doi.org/10.4103/0019-5154.190118

12. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, Chiller TM, Espitia V, Imtiaz R, Sosin D, Stroup DF, Tauxe RV, Vijayaraghavan M, Trostle M. Public health surveillance: A tool for targeting and monitoring interventions. Disease Control Priorities in Developing Countries 2nd edition [Internet]. International Bank for Reconstruction and Development/The World Bank; 2006 [cited 2021 Dec 29]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK11770/ PMID: 21250345

13. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. PLoS Comput Biol [Internet]. Public Library of Science (PLoS); 2013 Oct 17;9(10):e1003256. Available from: http://dx.doi.org/10.1371/journal.pcbi.1003256 PMCID: PMC3798275

14. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science [Internet]. American Association for the Advancement of Science (AAAS); 2014 Mar 14;343(6176):1203–1205. Available from: http://dx.doi.org/10.1126/science.1248506 PMID: 24626916

15. Pew Research. Demographics of Internet and Home Broadband Usage in the United States [Internet]. Pew Research Center: Internet, Science & Tech. 2021 [cited 2021 May 3]. Available from: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/

16. Wolfe DA. Ranked Set Sampling: Its Relevance and Impact on Statistical Inference. ISRN Probability and Statistics [Internet]. 2012 Dec 31;2012:1–32. Available from: http://dx.doi.org/10.5402/2012/568385

17. Culotta A, Ravi NK, Cutler J. Predicting the Demographics of Twitter Users from Website Traffic Data.

18. Nguyen D, Trieschnigg D, Doğruöz AS, Gravel R, Theune M, Meder T, de Jong F. Why Gender and Age Prediction from Tweets is Hard: Lessons from a Crowdsourcing Experiment. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers [Internet]. Dublin, Ireland: Dublin City University and Association for Computational Linguistics; 2014 [cited 2021 May 15]. p. 1950–1961. Available from: https://www.aclweb.org/anthology/C14-1184

19. Nearly Half of the Twitter Accounts Discussing "Reopening America" May Be Bots [Internet]. 2020 [cited 2021 May 15]. Available from: https://www.scs.cmu.edu/news/nearly-half-twitter-accounts-discussing-reopening-america-may-be-bots

20. González-Bailón S, Wang N, Rivero A, Borge-Holthoefer J, Moreno Y. Assessing the bias in samples of large online networks. Social Networks [Internet]. 2014;38:16–27. Available from: http://dx.doi.org/10.1016/j.socnet.2014.01.004

21. Twitter. Rate limits: Standard v1.1 [Internet]. 2021 [cited 2021 Mar 5]. Available from: https://developer.twitter.com/en/docs/twitter-api/v1/rate-limits

22. Zhao Y, Zhang H, Huo J, Guo Y, Wu Y, Prosperi M, Bian J. Mining Twitter to Assess the Determinants of Health Behavior towards Palliative Care in the United States. AMIA Summits on Translational Science Proceedings. 2020;2020:730.

23. Modave F, Zhao Y, Krieger J, He Z, Guo Y, Huo J, Prosperi M, Bian J. Understanding Perceptions and Attitudes in Breast Cancer Discussions on Twitter. arXiv:190512469 [cs, stat] [Internet]. 2019 May 22 [cited 2021 May 3]; Available from: http://arxiv.org/abs/1905.12469

24. Wang Y, Zhao Y, Bian J, Zhang R. Detecting Signals of Associations between Dietary Supplement Use and Mental Disorders from Twitter. 2018 IEEE Int Conf Healthc Inform Workshop (2018) [Internet]. 2018 Jun;2018:53–54. Available from: http://dx.doi.org/10.1109/ICHI-W.2018.00016 PMID: 31452863

25. Denny MJ, Spirling A. Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It. Polit Anal [Internet]. 2018 Apr;26(2):168–189. Available from: http://dx.doi.org/10.1017/pan.2017.44

26. Hargittai E. Is bigger always better? Potential biases of big data derived from social network sites. Ann Am Acad Pol Soc Sci [Internet]. SAGE Publications; 2015 May;659(1):63–76. Available from: http://dx.doi.org/10.1177/0002716215570866

27. Chiolero A, Santschi V, Paccaud F. Public health surveillance with electronic medical records: at risk of surveillance bias and overdiagnosis. Eur J Public Health [Internet]. Oxford University Press (OUP); 2013 Jun;23(3):350–351. Available from: http://dx.doi.org/10.1093/eurpub/ckt044 PMID: 23599219

28. Olteanu A, Castillo C, Diaz F, Kıcıman E. Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. Front Big Data [Internet]. 2019 Jul 11;2:13. Available from: http://dx.doi.org/10.3389/fdata.2019.00013

29. Audeh B, Calvier F-E, Bellet F, Beyens M-N, Pariente A, Lillo-Le Louet A, Bousquet C. Pharmacology and social media: Potentials and biases of web forums for drug mention analysis—case study of France. Health Informatics J [Internet]. 2020 Jun;26(2):1253–1272. Available from: http://dx.doi.org/10.1177/1460458219865128

30. Elkin LE, Pullon SRH, Stubbe MH. 'Should I vaccinate my child?' comparing the displayed stances of vaccine information retrieved from Google, Facebook and YouTube. Vaccine [Internet]. 2020 Mar;38(13):2771–2778. Available from: http://dx.doi.org/10.1016/j.vaccine.2020.02.041

31. Jaidka K, Giorgi S, Schwartz HA, Kern ML, Ungar LH, Eichstaedt JC. Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. Proc Natl Acad Sci USA [Internet]. 2020 May 12;117(19):10165–10171. Available from: http://dx.doi.org/10.1073/pnas.1906364117

32. McCosker A, Gerrard Y. Hashtagging depression on Instagram: Towards a more inclusive mental health research methodology. New Media Soc [Internet]. 2020 May 16;(23:7):1899–1919. Available from: http://dx.doi.org/10.1177/1461444820921349

33. Yang J-A, Tsou M-H, Jung C-T, Allen C, Spitzberg BH, Gawron JM, Han S-Y. Social media analytics and research testbed (SMART): Exploring spatiotemporal patterns of human dynamics with geo-targeted social media messages. Big Data & Society [Internet]. 2016;3(1):2053951716652914. Available from: http://dx.doi.org/10.1177/2053951716652914

34. Aslam AA, Tsou M-H, Spitzberg BH, An L, Gawron JM, Gupta DK, Peddecord KM, Nagel AC, Allen C, Yang J-A, Lindsay S. The Reliability of Tweets as a Supplementary Method of Seasonal Influenza Surveillance. J Med Internet Res [Internet]. 2014 Nov 14;16(11):e250. Available from: http://dx.doi.org/10.2196/jmir.3532

35. Jayawardhana UK, Gorsevski PV. An ontology-based framework for extracting spatio-temporal influenza data using Twitter. International Journal of Digital Earth [Internet]. 2019 Jan 2;12(1):2–24. Available from: http://dx.doi.org/10.1080/17538947.2017.1411535

36. Shan S, Yan Q, Wei Y. Infectious or Recovered? Optimizing the Infectious Disease Detection Process for Epidemic Control and Prevention Based on Social Media. INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH [Internet]. 2020 Sep;17(18). Available from: http://dx.doi.org/10.3390/ijerph17186853

37. Massey D, Huang C, Lu Y, Cohen A, Oren Y, Moed T, Matzner P, Mahajan S, Caraballo C, Kumar N, Xue Y, Ding Q, Dreyer R, Roy B, Krumholz H. Engagement With COVID-19 Public Health Measures in the United States: A Cross-sectional Social Media Analysis from June to November 2020. J Med Internet Res [Internet]. 2021 Jun 21;23(6):e26655. Available from: http://dx.doi.org/10.2196/26655

38. Effland T, Lawson A, Balter S, Devinney K, Reddy V, Waechter H, Gravano L, Hsu D. Discovering foodborne illness in online restaurant reviews. Journal of the American Medical Informatics Association [Internet]. 2018 Dec 1;25(12):1586–1592. Available from: http://dx.doi.org/10.1093/jamia/ocx093

39. Culotta A. Lightweight methods to estimate influenza rates and alcohol sales volume from Twitter messages. Lang Resources & Evaluation [Internet]. 2013 Mar;47(1):217–238. Available from: http://dx.doi.org/10.1007/s10579-012-9185-0

40. Alessa A, Faezipour M. Flu Outbreak Prediction Using Twitter Posts Classification and Linear Regression With Historical Centers for Disease Control and Prevention Reports: Prediction Framework Study. JMIR Public Health Surveill [Internet]. 2019 Jun 25;5(2):e12383. Available from: http://dx.doi.org/10.2196/12383

41. Margus C, Brown N, Hertelendy AJ, Safferman MR, Hart A, Ciottone GR. Emergency Physician Twitter Use in the COVID-19 Pandemic as a Potential Predictor of Impending

Surge: Retrospective Observational Study. J Med Internet Res [Internet]. 2021 Jul 14;23(7):e28615. Available from: http://dx.doi.org/10.2196/28615

42. Weeg C, Schwartz HA, Hill S, Merchant RM, Arango C, Ungar L. Using Twitter to Measure Public Discussion of Diseases: A Case Study. JMIR Public Health Surveill [Internet]. 2015 Jun 26;1(1):e6. Available from: http://dx.doi.org/10.2196/publichealth.3953

43. Tufts C, Polsky D, Volpp KG, Groeneveld PW, Ungar L, Merchant RM, Pelullo AP. Characterizing Tweet Volume and Content About Common Health Conditions Across Pennsylvania: Retrospective Analysis. JMIR Public Health Surveill [Internet]. 2018 Dec 6;4(4):e10834. Available from: http://dx.doi.org/10.2196/10834

44. Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, Conway M. Understanding Depressive Symptoms and Psychosocial Stressors on Twitter: A Corpus-Based Study. J Med Internet Res [Internet]. 2017 Feb 28;19(2):e48. Available from: http://dx.doi.org/10.2196/jmir.6895

45. Gattepaille LM, Hedfors Vidlin S, Bergvall T, Pierce CE, Ellenius J. Prospective Evaluation of Adverse Event Recognition Systems in Twitter: Results from the Web-RADR Project. Drug Saf [Internet]. 2020 Aug;43(8):797–808. Available from: http://dx.doi.org/10.1007/s40264-020-00942-3

46. Tacheva Z, Ivanov A. Exploring the Association Between the "Big Five" Personality Traits and Fatal Opioid Overdose: County-Level Empirical Analysis. JMIR MENTAL HEALTH [Internet]. 2021 Mar 8;8(3). Available from: http://dx.doi.org/10.2196/24939

47. Chary M, Genes N, Giraud-Carrier C, Hanson C, Nelson LS, Manini AF. Epidemiology from Tweets: Estimating Misuse of Prescription Opioids in the USA from Social Media. J Med Toxicol [Internet]. 2017 Dec;13(4):278–286. Available from: http://dx.doi.org/10.1007/s13181-017-0625-5

48. Iacus SM, Porro G, Salini S, Siletti E. An Italian Composite Subjective Well-Being Index: The Voice of Twitter Users from 2012 to 2017. Soc Indic Res [Internet]. 2020 Mar 16 [cited 2021 May 6]; Available from: http://link.springer.com/10.1007/s11205-020-02319-6

49. Delgado-Rodriguez M. Bias. Journal of Epidemiology & Community Health [Internet]. 2004 Aug 1;58(8):635–641. Available from: http://dx.doi.org/10.1136/jech.2003.008466

50. Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR, Chan A-W, Churchill R, Deeks JJ, Hróbjartsson A, Kirkham J, Jüni P, Loke YK, Pigott TD, Ramsay CR, Regidor D, Rothstein HR, Sandhu L, Santaguida PL, Schünemann HJ, Shea B, Shrier I, Tugwell P, Turner L, Valentine JC, Waddington H, Waters E, Wells GA, Whiting PF, Higgins JPT. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ [Internet]. 2016 Oct 12;i4919. Available from: http://dx.doi.org/10.1136/bmj.i4919

51. Centers for Disease Control and Prevention. United States COVID-19 Cases and Deaths by State over Time [Internet]. 2021 [cited 2021 Dec 31]. Available from:

https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36

52. Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report (MMWR) | MMWR [Internet]. 2021 [cited 2021 Dec 31]. Available from: https://www.cdc.gov/mmwr/index.html

53. Experian. Experian Marketing Services [Internet]. 2021 [cited 2021 Jul 14]. Available from: https://www.experian.com/marketing-services/

54. Peggy Baker, Jonaki Bose, Joseph Gfroerer, Beth Han, Sarra L. Hedden, Arthur Hughes, Michael Jones, Joel Kennet. Results from the 2010 National Survey on Drug Use and Health: Summary Of National Findings 2011. Center for Behavioral Health Statistics and Quality; 2011 Sep.

55. Ohio Department of Health. Ohio Department of Health [Internet]. 2021 [cited 2021 Dec 31]. Available from: https://odh.ohio.gov/wps/portal/gov/odh/home

56. Sharecare, Inc. Community Well-Being Index [Internet]. 2021 [cited 2021 Dec 31]. Available from: https://wellbeingindex.sharecare.com/

57. ISTAT. La soddisfazione dei cittadini per le condizioni di vita [Internet]. 2017 [cited 2021 Jul 19]. Available from: https://www.istat.it/it/files//2018/01/Soddisfazione-cittadini.pdf

58. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases [Internet]. 2020 May;20(5):533–534. Available from: http://dx.doi.org/10.1016/S1473-3099(20)30120-1

59. Cneters for Disease Control and Prevention. Multiple Cause of Death, 1999-2019 Request [Internet]. 2021 [cited 2021 Dec 31]. Available from: https://wonder.cdc.gov/mcd-icd10.html

60. CDC. U.S. Outpatient Influenza- like Illness Surveillance Network (ILINet): Percentage of Visit for ILI by Age Group [Internet]. 2020 [cited 2021 Jul 21]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/10232020/percent-ili-visits-by-age.html

61. CDC. National, Regional, and State Level Outpatient Illness and Viral Surveillance [Internet]. 2021 [cited 2021 Jul 14]. Available from: https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html

62. Cesare N, Grant C, Nguyen Q, Lee H, Nsoesie EO. How well can machine learning predict demographics of social media users? arXiv:170201807 [cs] [Internet]. 2018 May 30 [cited 2021 Jul 21]; Available from: http://arxiv.org/abs/1702.01807

63. Preoţiuc-Pietro D, Volkova S, Lampos V, Bachrach Y, Aletras N. Studying User Income through Language, Behaviour and Affect in Social Media. Braunstein LA, editor. PLoS

ONE [Internet]. 2015 Sep 22;10(9):e0138717. Available from: http://dx.doi.org/10.1371/journal.pone.0138717

64. Alabdulkreem E. Prediction of depressed Arab women using their tweets. JOURNAL OF DECISION SYSTEMS [Internet]. Available from: http://dx.doi.org/10.1080/12460125.2020.1859745

65. Zhao Y, Guo Y, He X, Wu Y, Yang X, Prosperi M, Jin Y, Bian J. Assessing mental health signals among sexual and gender minorities using Twitter data. Health Informatics J [Internet]. 2020 Jun;26(2):765–786. Available from: http://dx.doi.org/10.1177/1460458219839621

66. Magdy W, Elsayed T. Adaptive Method for Following Dynamic Topics on Twitter. ICWSM. 2014.

67. Morstatter F, Pfeffer J, Liu H, Carley KM. Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. arXiv:13065204 [physics] [Internet]. 2013 Jun 21 [cited 2021 May 6]; Available from: http://arxiv.org/abs/1306.5204

68. Pfeffer J, Mayer K, Morstatter F. Tampering with Twitter's Sample API. EPJ Data Sci [Internet]. 2018 Dec;7(1):1–21. Available from: http://dx.doi.org/10.1140/epjds/s13688-018-0178-0

69. Hellström T, Dignum V, Bensch S. Bias in Machine Learning -- What is it Good for? arXiv:200400686 [cs] [Internet]. 2020 Sep 20 [cited 2021 Jul 20]; Available from: http://arxiv.org/abs/2004.00686

70. Battaglia E, Bioglio L, Pensa RG. Towards content sensitivity analysis. Lecture Notes in Computer Science [Internet]. Cham: Springer International Publishing; 2020. p. 67–79. Available from: http://dx.doi.org/10.1007/978-3-030-44584-3_6