Review

# Methods for Social Media Monitoring Related to Vaccination: Systematic Scoping Review

Emilie Karafillakis[1*], MSc; Sam Martin[1*], PhD; Clarissa Simas[1], MSc; Kate Olsson[2], MPH; Judit Takacs[2,3], PhD; Sara Dada[1], MSc; Heidi Jane Larson[1,4], PhD

[1]London School of Hygiene & Tropical Medicine, Vaccine Confidence Project, London, United Kingdom

[2]European Centre for Disease Prevention and Control, Stockhom, Sweden

[3]Centre for Social Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[4]Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, United States

[*]these authors contributed equally

**Corresponding Author:**
Emilie Karafillakis, MSc
London School of Hygiene & Tropical Medicine
Vaccine Confidence Project
Keppel Street
London, WC1E 7HT
United Kingdom
Phone: 44 2079272831 ext 2831
Email: emilie.karafillakis@lshtm.ac.uk

## *Abstract*

**Background:** Social media has changed the communication landscape, exposing individuals to an ever-growing amount of information while also allowing them to create and share content. Although vaccine skepticism is not new, social media has amplified public concerns and facilitated their spread globally. Multiple studies have been conducted to monitor vaccination discussions on social media. However, there is currently insufficient evidence on the best methods to perform social media monitoring.

**Objective:** The aim of this study was to identify the methods most commonly used for monitoring vaccination-related topics on different social media platforms, along with their effectiveness and limitations.

**Methods:** A systematic scoping review was conducted by applying a comprehensive search strategy to multiple databases in December 2018. The articles' titles, abstracts, and full texts were screened by two reviewers using inclusion and exclusion criteria. After data extraction, a descriptive analysis was performed to summarize the methods used to monitor and analyze social media, including data extraction tools; ethical considerations; search strategies; periods monitored; geolocalization of content; and sentiments, content, and reach analyses.

**Results:** This review identified 86 articles on social media monitoring of vaccination, most of which were published after 2015. Although 35 out of the 86 studies used manual browser search tools to collect data from social media, this was time-consuming and only allowed for the analysis of small samples compared to social media application program interfaces or automated monitoring tools. Although simple search strategies were considered less precise, only 10 out of the 86 studies used comprehensive lists of keywords (eg, with hashtags or words related to specific events or concerns). Partly due to privacy settings, geolocalization of data was extremely difficult to obtain, limiting the possibility of performing country-specific analyses. Finally, 20 out of the 86 studies performed trend or content analyses, whereas most of the studies (70%, 60/86) analyzed sentiments toward vaccination. Automated sentiment analyses, performed using leverage, supervised machine learning, or automated software, were fast and provided strong and accurate results. Most studies focused on negative (n=33) and positive (n=31) sentiments toward vaccination, and may have failed to capture the nuances and complexity of emotions around vaccination. Finally, 49 out of the 86 studies determined the reach of social media posts by looking at numbers of followers and engagement (eg, retweets, shares, likes).

**Conclusions:** Social media monitoring still constitutes a new means to research and understand public sentiments around vaccination. A wide range of methods are currently used by researchers. Future research should focus on evaluating these methods to offer more evidence and support the development of social media monitoring as a valuable research design.

XSL•FO
**RenderX**

**KEYWORDS**

vaccination; antivaccination movement; vaccination refusal; social media; internet; research design; review; media monitoring; social listening; infodemiology; infoveillance

## Introduction

Although public questioning of vaccination is as old as vaccination itself [1], continuous advancements in the global communication landscape, associated with the rise of social media as an interactive health information ecosystem, have contributed to the unmediated spread of vaccine hesitancy [2]. This new boundless information ecosystem has shaped the nature of conversations about vaccination, with evidence showing that social media can facilitate the quick diffusion of negative sentiments and misinformation about vaccination [2-6]. Furthermore, individuals have been found to more commonly engage with negative information around vaccination than positive content [7-10]. In this context, public trust in information provided by authorities and experts can decrease [11-13], influencing vaccine decisions [14]. Recent evidence has shown that social media users tend to cluster and create so-called "echo chambers" based on their views toward vaccination [15]; however, Leask et al [16] highlight that "a patient's trust in the source of information may be more important than what is in the information," stressing the importance of reaching individuals, across all clusters, through trustworthy sources.

Social media monitoring (infoveillance) provides opportunities to listen, in real time, to online narratives about vaccines, and to detect changes in sentiments and confidence early [17]. Information gathered from social media monitoring is crucial to inform the development of targeted and audience-focused communication strategies to maintain or rebuild trust in vaccination [17,18]. However, as social media monitoring can be resource- and time-intensive, and can raise issues of confidentiality, transparency, and privacy [17,19,20], evidence of public health communities investing in such listening mechanisms remains sparse.

The aim of this scoping review was to systematically summarize the methodologies that have been used to monitor and analyze social media on vaccination using an innovative three-step model. The findings presented in this paper come from a broader European Centre for Disease Prevention and Control (ECDC) technical report [21]. The aim of the ECDC report was to provide guidance for public health agencies to monitor and engage with social media, whereas this paper primarily focuses on the academic implications of social media monitoring. The specific objectives of this scoping review were to (1) identify the methods most commonly used for monitoring different social media platforms; and (2) identify the extent to which methods have been evaluated, along with their effectiveness and limitations.

## Methods

### Design

Systematic scoping reviews are used to map international literature with the aim of clarifying "working definitions and conceptual boundaries of a topic or field" [22] as well as identifying how research is conducted [22-24]. Systematic scoping reviews focus on scoping larger, more complex, and heterogeneous topics than systematic literature reviews. A systematic scoping review approach was therefore adapted to fulfill the goal of summarizing study methodologies used to monitor social media content around vaccination. The methodology for this scoping review was based on the work of Arksey et al [23] and Peters et al [24].

### Framing Social Media

Kaplan et al [25] define social media as "a group of internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content." They further classify social media into blogs, collaborative projects (eg, Wikipedia), social networking sites (eg, Facebook), content communities (eg, YouTube), virtual social worlds (eg, Second Life), and virtual game worlds (eg, World of Warcraft) [25].

However, social media is not merely an information tool but also represents a continuously evolving social environment directly influenced by how individuals produce and share content, and interact with each other. For the purpose of this scoping review, we consider social media as not simply a means of communication but further a space within which individuals socialize and organize. This review therefore focuses on social networking sites and content communities, and excludes online platforms that do not have social interactions as their main purpose (eg, blogs or websites with a comments section).

### Search Strategy and Screening Process

The search strategy for the scoping review was developed by librarians at ECDC and researchers at the Vaccine Confidence Project (VCP), and was peer-reviewed to balance feasibility and comprehensiveness, including both social media and vaccination-related English keywords (see Multimedia Appendix 1). The search was conducted by one VCP researcher on the EMBASE database, and was adapted to search the PubMed, Scopus, MEDLINE, PsycINFO, PubPsych, Open Grey, and Web of Science databases in December 2018.

Identified articles were exported into Endnote, and duplicates were removed based on ECDC guidelines consisting of 6 rounds of deduplication looking for articles with similar author, year, and title; title, volume, and pages; author, volume, and pages; year, volume, issue, and pages; title; and author and year. The automated deduplication function in Endnote was not used, and articles were compared visually to ensure that only true

duplicates were removed. Two VCP reviewers independently screened articles by title and abstract and by full text using a set of predefined inclusion and exclusion criteria. Disagreements were resolved by discussion.

Articles were included if they described studies performed to monitor or analyze data collected from social media around vaccination. The definition of social media described above was used as one of the inclusion criteria, limiting results to social networking sites and content communities. No restrictions were made with respect to location or language, as a team of official translators was available at the ECDC.

Articles were excluded if they were published before 2000 or if they were not about human vaccines. Articles that monitored online media (eg, news, websites) but did not collect any data from social media were excluded. The following article types were also excluded: conference abstracts, editorials, commentaries, and letters to the editor.

## Data Management and Analysis

Two VCP researchers extracted the following data from the included articles: country, aim, study population, period of monitoring, vaccine, social media, media monitoring methodologies (tool for data collection, keywords, exclusion criteria, geolocation), analysis (sentiment coding and analysis,

reach, spread and interaction analyses, other types of analyses), results (number of posts), and evaluation and limitations.

To facilitate the description of social media monitoring methods, a three-step model of social media monitoring was developed (Figure 1), including (1) preparation, (2) data extraction, and (3) data analysis steps. The preparation phase consists of defining the purpose of social media monitoring and addressing any ethical considerations. The data extraction phase includes selecting data extraction tools and periods of monitoring, developing comprehensive search strategies, and extracting the data. Finally, the data analysis stage includes geolocation, trends, content, sentiments, and reach analyses. The findings summarized in this paper are organized according to this three-step model.

Three researchers summarized, charted, and analyzed the data. A descriptive analysis was conducted for the types of data collection tools used to gather data from social media, the keywords and search strategies used, and the various analytical methods. These researchers reviewed and compared results in the data extraction sheet, listed and identified the frequency of different methods used for social media monitoring, and identified common themes. Two researchers met to discuss the findings and interpret them together with contextual information, identifying needs for further research.

**Figure 1.** The three-step model of social media monitoring.

## *Results*

### Included Studies

The search strategy generated 15,435 articles, from which 7539 duplicates and 7628 irrelevant articles were excluded after screening by title and abstract (Figure 2). From the 268 articles screened by full text, 182 were excluded for the following reasons: not about social media (n=141); no data provided (n=19); conference abstracts, editorials, or letters to the editor (n=6); article not accessible (even after enquiring multiple libraries and contacting authors) (n=4); article containing data already published in another included article (n=1); and not on vaccination (n=1). At the end of the screening process, 86 articles in English, Spanish, and Italian were included for analysis. Articles in Spanish and Italian were analyzed by a researcher fluent in these two languages.

**Figure 2.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.
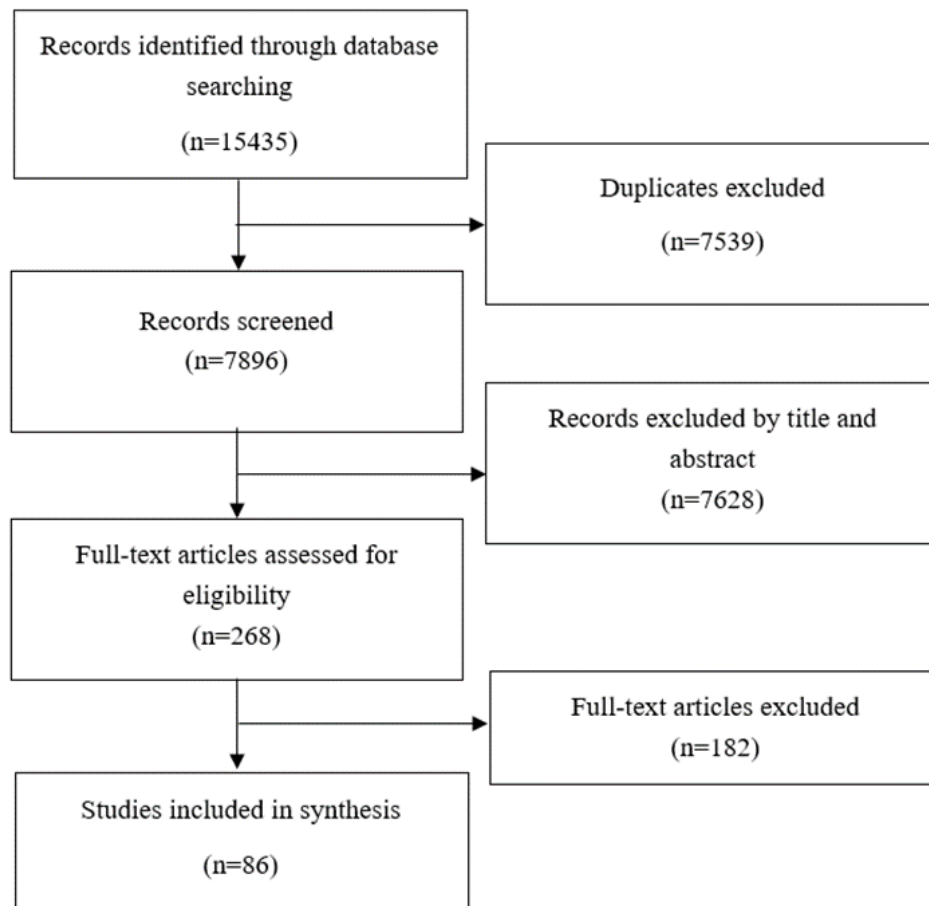


### Study Characteristics

The first study identified on social media monitoring around vaccination was published in 2006 [26], with an increasing number of studies published yearly since then. Most studies analyzed online discourse on Twitter (n=42) [8,27-67], YouTube (n=12) [68-79], Facebook (n=11) [80-90], and online forums (eg, babytree, mothering.com, mumsnet, KaksPlus; n=9) [26,91-98]. A diversification of social media platforms can be observed in recent years, with studies of platforms such as Pinterest (n=1) [99], Weibo (n=1) [100], Reddit (n=1) [101], or Yahoo! Answers (n=2) [102,103] all published after 2015. Seven studies monitored a mix of social media platforms [104-110]. More detailed study characteristics are provided in Multimedia Appendix 2.

### Social Media Monitoring Methods

A range of methods were used across the 86 included studies to monitor social media, most of which have not been evaluated in terms of their accuracy and reliability. In the following sections, these methods are described based on the three steps of social media monitoring proposed in this paper: preparation, data extraction, and data analysis.

### Preparation Phase

#### *Defining the Purpose of Social Media Monitoring*

The main objective of the majority of studies in this review (55/86, 64%) was to better understand how vaccination is portrayed on social media, whether through the analysis of online discourse or sentiments, or by looking at how information is produced, shared, and engaged with [8,26,27,30,33, 39,42-44,46-51,53-56,58-62,68-80,82-88,90-94,98,99,101,102,105,106].

Many studies (15/86, 17%) used social media monitoring as a way to better understand general public discussions on vaccination, assuming that online discussions are a good proxy for vaccine confidence in a country or region [29,32,34,36-38,45,57,63,65,67,81,95,96,103]. This comes with important limitations due to the lack of representativeness of social media populations. Individuals discussing vaccination on social media tend to come from specific population groups, usually younger or female groups [61]. Furthermore, data extracted from social media platforms are often not representative of the entire online discourse around vaccination on these platforms (see Data Extraction Phase section).

Other objectives included estimating correlations between online activity around vaccination and coverage or outbreak data [40,41,66,104,107,109], describing systems for monitoring vaccination online [28,52,64,97], investigating relations between news media and social media posts [100,108,110], understanding the contribution of bots or trolls to online content about vaccination [31], examining political references to vaccination on social media [35], and detecting anxiety-related adverse events following immunization [89].

### Addressing Ethical Considerations

Access to social media data is becoming increasingly restricted, as some users set their profiles, conversations, or pages as "private" [111]. Yet, only 15% (13/86) of the studies included in this review were found to have been reviewed and to have received approval from an institutional ethics review board for their study [32,36-38,50,56,67,81,83,85,86,90,93]. An additional 9 studies sent protocols to institutional or ethical review boards and were deemed exempt because they only analyzed public data and social media users were not considered as "human research subjects" [8,31,44,48,60,69,70,75,80]. One study also explained that guidelines from an institutional review board were considered and applied during the study to protect social media users [92].

However, even once ethical approval had been obtained, questions about anonymity, confidentiality, and informed consent remained. For example, one study explained that anonymization of data is extremely difficult to maintain on social media, as content and quotes (whether from private or public data) can easily be traced back to users, revealing their identity [54]. The authors of another study performed on Facebook in Israel explained that although they anonymized their data, informed consent was not required as "subjects would expect to be observed by strangers" when posting messages on the internet [85]. Some studies also discussed the limitations of focusing on public data and the distorted view this creates [28,43,54,61,83,85,88,97,109].

### Data Extraction Phase

#### Selecting Data Extraction Tools and Periods of Monitoring

Studies included in this review were found to use different monitoring tools to extract data from social media platforms. Thirty-five studies used manual browser search tools such as search bars available on Twitter, YouTube, or Facebook [26,60,68-81,83,85,86,89-100,102,104,106,109]. Due to browser and user interface limitations, studies that used manual browser search functions were time-consuming and collected small amounts of data over short periods of time. In one study, the analysis had to be limited to 30 Facebook pages [80], which affected the possibility of capturing data over different periods (ie, possibly missing trends in the number of posts around influenza seasons). Furthermore, owing to the time needed to assess Facebook pages, it was found to be impractical to analyze each complete page in detail. Another limitation comes from browser cookies and personal tracking algorithms imposed by search engines, which can influence manual search results and their listed popularity. Researchers from another study indicated that findings from both the Google and Facebook searches were dependent on the geographic location of the reviewer's browser settings [89]. Researchers using manual search engines are also restricted to the way search results are presented on different platforms; for example, Pinterest does not list its pins chronologically and does not provide exact time stamps [99]. This made using a more conventional content analysis sampling method (eg, a constructed 2-week time period) virtually impossible in this particular study.

Forty-nine studies used either social media application program interfaces (APIs) (n=24) [29,31-33,35,37-41,43,45,48,54,56-58,65-67,82,84,87,88,101,103], automated monitoring tools (n=20) [27,36,42,44,46,47,49-53,55,59,61,62,64,105,107,108,110], or a combination of both (n=5) [8,28,32,34,63] to extract data from social media platforms. The term API refers to a software intermediary that allows two apps to talk to each other [112]. APIs pull and interpret data from servers storing information for Facebook, Twitter, YouTube, Reddit, and many more platforms. It is important to note that APIs do not provide comprehensive access to all social media content, and often only pull random samples of content; for example, Twitter provides access to roughly 1% of public Tweets through its API [113]. Automatic monitoring tools refer to automated web platforms that access social media data via APIs. These automated tools come with user-friendly interfaces, which can be free (with limited access to a random sample of all posts), open source (open to development from other developers), or commercial (where access to a larger percentage of posts is allowed, which can be real-time or archival via a subscription pricing structure). Regardless of the data collection period, studies with the highest number of results and the most robust datasets consistently came from the use of social media APIs or automatic data sampling. The "Yahoo! Answers" API provided the largest sample size from a single platform over a sampling period of 5 years (16 million messages) [103] and Crimson Hexagon was the automated platform that provided the largest mixed sample size, with a mixture of 58,078 Facebook posts and 82,993 tweets over a 7.5-year period [105]. The Yahoo! Answers study found that the API data were difficult to stratify by age, gender, income, education level, or marital status, which may have limited generalizability [103]. Similarly, Smith et al [57] found that using the Twitter API to use social media discussions as a proxy for the population at large is problematic. The difficulty in finding the correct self-assigned demographic of users, and whether they are real users or automated bots, makes the findings less generalizable. Finally, a large number (n=45) of studies using either APIs or

XSL•FO

RenderX

automated software focused on Twitter due to the ease of access given by the platform to its data stream compared to other platforms, which may give a skewed perspective of social media attitudes toward vaccines [8,27-59,61-67,105,107,108,110].

Although using a mixture of tools to collect data from social media is possible, only one study used a combination of APIs and manual tools [30].

### Developing Comprehensive Search Strategies

This review found a diverse range of search strategies developed to extract data from social media platforms. Simple search strategies with one to three keywords were most common. Only 10 studies used more extensive search strategies with Boolean operators to link keywords (eg, AND, OR) or truncations to identify words with different endings (eg, vaccin*) [27,29,42,43,50,51,66,101,105,108]. Although simpler search strategies were perceived as a limitation by some [8,71,73], no data were available on the accuracy of short strategies as opposed to longer and more complex search strategies. Studies that evaluated their search strategies found that the categorization of keywords into "relevant," "semirelevant," and "nonrelevant" can increase precision [43], and that keywords should reflect cultural and normative differences [108].

Across all studies, most keywords were related to vaccines (ie, synonyms of the word "vaccine" or brand names of vaccines) and vaccine-preventable diseases. Some studies also searched for adverse events claimed to be linked to vaccination by the public (eg, autism, autoimmune disorders), keywords related to specific controversies (eg, mercury, big pharma, aluminum), or the names of people involved in controversies (eg, Jenny McCarthy, Andrew Wakefield). In addition to keywords, certain studies used hashtags (eg, #vaccine, #cdcwhistleblower, #vaccineswork) [8,30,34,44,48,49,55,60,61,105], questions inputted into search engines (eg, "should I get the HPV [human papillomavirus] vaccine?") [75,102], or phrases to refer to specific events (eg, "fainting in school children after vaccine") [89].

Predefined exclusion criteria were also used to screen data and exclude irrelevant or duplicate results. The question of how to deal with data from dubious sources, trolls, or bots was raised, and although researchers in two studies decided to exclude them, two studies specifically analyzed them and acknowledged their impact on the quality and validity of their findings [31,45,57,108].

## Data Analysis Phase

### Analyzing Metadata, Including Geolocation

The included studies analyzed a range of metadata, from the number of posts to users' characteristics. Information about the geographical source of social media data was extremely difficult to obtain, as this information was often private, not provided by social media users, or, as one study conducted on Twitter explained, because "accurate location information can be found in only a small proportion of tweets that have coordinates stored in the metadata of the tweet" [40]. This could explain why most of the studies in this review were performed "globally" (n=41).

Despite these challenges, three types of strategies were used to restrict data to certain regions or countries: using keywords in local languages; using location-specific search terms (eg, United Kingdom, Scotland); and directly identifying local or national Facebook groups, pages, or online discussion forums [50,82,86,90,91,93]. Once social media posts were collected, other tools were used to identify and analyze the precise location of data. Some studies manually screened content or collected metadata [39,56,108], whereas others used automated mechanisms and software (Carmen, Geodict, Nominatim, GeoSocial Gauge) to retrieve this information from Twitter [28,29,40,41,54,61,66]. Two studies used dictionaries of terms for geographical entities of countries (GeoNames and the US Office of Management and Budget's Metropolitan and Micropolitan Statistical Areas) to automatically identify mentions of countries or cities in social media posts or profile pages [29,61]. Some authors also explained that most bots spreading negative content about vaccination online do not report their locations, which could explain why most tweets with geolocation information available were more positive toward vaccination [61].

### Analyzing Trends, Content, and Sentiments

Once data from social media were extracted from studies, different analyses were performed, ranging from detecting the number of posts available over a period of time to more detailed content analysis to identify the frequency of particular concerns or conspiracies around vaccination [8,27,46-48,53,59,60,68-71,75-77,79,82,83,90,94,99]. Several studies also performed qualitative thematic analysis [45,56,71,83,85,86,95,99,110], or language and discourse analysis [26,50,52,84,92,93,106]. Four studies compared social media posts to disease incidence or outbreak cases [40,103,105,107].

The most common type of analysis looked at sentiments expressed toward vaccination (70%, 60/86). Sentiments can be understood in a variety of ways, reflected by the range of words identified to designate sentiments toward vaccination across all studies. Most studies used the terms "negative" (n=33), "positive" (n=31), or "neutral" (n=37); however, each study defined these in a slightly different way, which could have influenced study findings and what was perceived by researchers as "negative" or "positive." Other common sentiments were anti- or provaccine, encouraging or discouraging, ambiguous, or hesitant. Only two studies provided a more comprehensive list of sentiments such as frustration, humor, sarcasm, concern, relief, or minimized risk [53,107]. One study also looked at sentiment as a "yes or no" question: "does this message indicate that someone received or intended to receive a flu vaccine?" [41]. In one study, the World Health Organization determinants of the vaccine hesitancy framework were used to design and test a list of sentiments [90].

Sentiment was determined not only by looking at social media posts or comments but also by coding links, headlines, sources, images, captions, or hashtags [27,35,42,46,60,64,99]. Coding hashtags was sometimes difficult; for example, those using the hashtag #antivaxxers were often denouncing vaccine hesitancy. Similarly, "positive" hashtags such as #provaxxers can be used in a negative context to criticize those who promote vaccination

[27]. Coding sarcasm, irony, slang, and hyperboles was also complicated and prone to subjectivity biases [27,32,34,42,44,63,101].

Sentiment analyses were performed manually (40/86, 67%) or using an automated system (19/86, 32%). When data were analyzed manually, studies used multiple coders (between 2 and 4) and assessed interrater reliability scores to ensure accuracy and reliability. Many studies also emphasized the need to provide coders with training and codebooks with precise definitions of codes [8,26,31,32,46,47,53,60,71,73,83,90, 107,109]. Manual sentiment analysis was prone to limitations, particularly because it relied on subjective coding and was labor-intensive, thereby reducing the total number of posts that could be analyzed by a single person [8,43,68,107].

Sentiments were also analyzed using automated systems, with most studies using such systems performed on Twitter (16/19, 84%). Leverage or supervised machine learning was used to code sentiments by training machines to learn how to code different sentiments using a set of manually coded results (ranging between 693 and 8261 posts) [28,34,37-39,41,48, 49,51,57,62,63,66,67,82]. An alternative to manually coding some results to train the machine was to use Amazon Mechanical Turk [41,51,57,66]. Other automated systems that have been used to code sentiments included Latent Dirichlet Allocation, an unsupervised machine-learning algorithm that automatically determines topics in a text [57,101]; Naive Bayes [65]; Lightside [61]; BrightView classifier from Crimson Hexagon [105]; and Topsy [29,44]. Using such programs also came with limitations, including the aptitude of machines to correctly detect sentiments around vaccination, the reliance on manual coding of some part of the data to train the system prone to researchers' biases and subjectivity, and the need for high computational and technical skills [28,44,101].

### Assessing the Reach of Social Media

Overall, 49 studies measured potential social media reach and thus estimations of the number of people that see content posted on social media. Reach was determined by the number of followers a user had, as well as the number of engagements with a post (eg, retweeted, shared, saved, liked, and commented upon).

Most studies provided a short descriptive summary of the reach of social media posts, whereas some proposed more detailed or comprehensive analyses. Interactions between different social network communities were studied to understand how information can spread and be shared on social media. Studies found that analyzing retweets was useful to understand the spread of certain sentiments toward vaccination and, in the case of disease outbreaks, to detect how the spread of social media information online can impact vaccination coverage. One particular study investigated how two kinds of communities interacted with each other within conversations about health and its relation to vaccines [62]. From a retweet network of 660,892 tweets published by 269,623 users, the study compared "structural community" with another "opinion group," and used community detection algorithms and autotagging to measure the interaction, sentiment, and influence that retweets had in conversations between the two communities [62]. Similarly,

another study focused on shared concerns about the HPV vaccine and assessed how international followers express similar concerns to those of the groups or individuals they follow [56].

Another study examined communication patterns revealed through retweeting, assessing the impact of various sources of information, contrasting diverse types of authoritative content (eg, health organizations and official news organizations) and grassroots campaign arguments (with the antivaccination community views serving as a prototypical example) [54]. Finally, one study looked at tweeted images, and evaluated predictive factors for determining whether an image was retweeted, including the sentiment of the image and the objects shown in the image [33].

## Discussion

### Principal Findings

Over 80 articles have been published on social media monitoring around vaccination. This growing academic interest, particularly since 2015, acknowledges the role of social media in influencing public confidence in vaccination, and emphasizes the need to better understand the types of information about vaccination circulating on social media and its spread within and between online social networks [114,115]. Social media monitoring still constitutes a relatively new research field, for which tools and approaches continue to evolve. A wide range of methods, varied in style and complexity, have been identified and summarized through this systematic scoping review.

In an effort to summarize media monitoring articles, we developed a three-step model for this review. The first stage, preparation, consists of defining the purpose of social media monitoring while considering any potential ethical issues. The second stage, data extraction, should include the selection of data extraction tools as well as periods of monitoring, and the development of targeted, comprehensive, and precise search strategies. Finally, the third stage, data analysis, could focus on different types of analyses: metadata and geolocation, trends, content, sentiment, or reach. The model was found to be useful in structuring methodologies for social media monitoring, and could be used in the future as a standardized protocol for performing social media monitoring. Further research could be performed to evaluate different components of the model, and propose a more detailed and complex framework for media monitoring.

### Standardization of Social Media Monitoring Methods

Although the large number of articles identified via the scoping review provided sufficient evidence to summarize methods that have been used to monitor social media, almost none of the articles evaluated the precision and accuracy of their monitoring and analysis methodologies. Furthermore, researchers have not drawn on a coherent body of agreed-upon methodologies, and instead created an amalgamation of methodological choices that sets no standards for the right sample size, no recommended time period for different types of analyses per platform, or no recommendations for studying the extremes of positive or negative views (which are not always representative of the general population) [116]. There is also a lack of standardization

of which specific API tools or analytical classifiers to use for analyzing social media discourse, interaction, or trends.

There have been recent calls to better standardize social media monitoring methodologies, including practices such as search strategies, so that the quality of the data query is reflected in more accurate and precise data and study findings [117]. However, it may be apt that social media monitoring remains a flexible research design, as the nature and access to social media discourse on vaccination is continuously evolving. The fast-evolving nature of different social media platforms, the crossover of shared data, boundaries to privacy, and public policy surrounding public discourse on vaccination and disease outbreaks may also necessitate a more methodologically diverse approach to keep up with ever-changing developments. Although standardization may not be the best practice for this relatively new research design, there is a need to evaluate the different tools that will be used at each stage of social media monitoring to determine which ones offer the most precise, accurate, and representative results.

## Establishing the Purpose of Social Media Monitoring

Using social media to understand prevailing issues of interest and concern in certain communities can be a useful listening tool for public health institutions, which can then use media monitoring to detect key themes or questions around vaccination circulating in the population. However, many studies included in this review explicitly discussed limitations regarding the lack of population representativeness in investigating social media content. Evidence shows that social media users often represent specific population groups in terms of age, gender, education level, or socioeconomic status [118]. For instance, users discussing vaccination online were found to be younger and female [61]. Another challenge comes from the fact that content being shared by social media users is not always representative of their personal views or feelings, with evidence showing that social media content is often more extreme or impulsive [119]. Due to these issues, social media monitoring is best seen as an alternative to surveys or qualitative interviews in obtaining data about vaccination beliefs and opinions, without assuming representativeness of total populations but rather specific interest groups.

Social media users could also be considered as a configuration of a research population group, and the field of social media monitoring could be seen as an opportunity to understand what information users are exposed to, and how information about vaccination is shared and spread online. In this way, social media monitoring would be used as a new research methodology to study a new type of population. Social media monitoring comes with representativeness challenges of its own, as access to data becomes limited due to inaccessible private content, the challenge of studying all social media platforms at once, or limitations imposed by automated software. However, social media monitoring opens the door to more dynamic research that continuously evolves and responds to a perpetually changing world.

## Important Ethical Considerations

The considerable increase in the number of social media monitoring studies poses questions regarding the safe use of data available online. Even though researchers in previous studies may not have been legally compelled to obtain ethics approval, the lack of guidance on good ethical conduct when using social media information is a cause for concern. Issues of confidentiality and anonymization of data still arise, as some studies included in this review published screenshots of users' data that included users' profile names. Another issue relates to data coming from minors, which should be considered more carefully, even when publicly available [120]. Although these concerns should not unduly hinder the development of social media monitoring, they should highlight the need for guidelines to ensure ethical conduct and respect for social media users, and the importance of submitting research proposals to ethics boards.

Recent controversies with regard to the exploitation of users' data in the Facebook and Cambridge Analytica scandal, and the public outcry of users feeling unnerved being monitored and manipulated, have indeed opened up conversations and legislation around the ethics of handling user data from social media in research [121]. The overall aim of the 2018 EU General Data Protection Regulation (GDPR) is to increase people's control over their own personal data and "to protect all EU citizens from privacy and data breaches in an increasingly data-driven world" [122]. For companies, organizations, and researchers, this means obtaining consent for using and retaining customers' personal data, while granting more rights to the "data subject" to be informed and to control how their personal data are used. Such legislation may change the way future researchers must anonymize data, as well as restrict what sections of social media platforms (eg, public vs private) are available for research [123].

## Accessing Data From Different Social Media Platforms: The Twitter Bias

The majority of studies that had the largest datasets, collected over longer time periods, were those with access to social media platforms' APIs or automated data collection tools. Studies that had smaller samples and used less sophisticated keyword searches were those that relied on manual data collection, and were thus constrained by time, resources, and the limitations of the browser tools used. Although some studies also discussed the time-consuming nature of manual data collection methodologies, the time required to perform searches was not commonly discussed. More research could be performed to evaluate the clear benefits and limitations of manual and automated data extraction tools, including the time required to complete searches. Studies that used the paid version of APIs via automated monitoring software seemed to have a more representative sample, as access to paid data offers access to all historical and current posts. However, there are still issues with the relative opaqueness of the paid access to Twitter, Facebook, or YouTube APIs, which do not advertise the mechanisms behind collection of data, do not inform researchers of what percentage of "all" data they are given, and may thus not provide representative data [124,125]. This also prevents researchers

from fully comparing studies over time, as the API sampling algorithm itself will change. It is arguable that such an environment presents risks and opportunities both for data collection strategies in terms of availability and data privacy issues, as well as an evolution in our fundamental understanding of how social media research fits into the rapidly changing public discourse in relation to vaccines. Finally, the financial cost associated with the use of most APIs and automated tools could constitute a barrier to those in lower resource settings [126].

One of the main reasons for the bias toward the use of Twitter in a majority of studies within this review may be because Twitter provides the most openly available API, both for free and with paid access [116]. However, studies using these freely collected tweets only have access to a small (1%) sample of all tweets, creating representativeness challenges [113]. Accessing the free Twitter API also raises issues around periodical collection due to restrictive access to intermittent collection points. This means that any data collection is limited to pockets of time that are not necessarily continuous, truncating the 1% sample into different time periods [127]. This focus on Twitter constitutes an important bias for social media monitoring research, as it fails to capture the real-time evolution of the social media environment, and the flow of users and content from platform to platform [128].

Finally, although subscription-based data analytics companies provide more comprehensive access to other APIs such as YouTube and Instagram, the data provided by these companies can be more skewed toward those of brand marketing (eg, brand strength, brand influencers, and brand trends) [129]. However, there have been growing opportunities in recent years to allow academics to work in partnership with data analytics companies to forge a better understanding of how to look at social media images and text from a social sciences and public policy perspective [130].

## Social Media Analyses: Complexity of Analyzing Sentiments

Although social media data can be analyzed in various ways, ranging from analyzing trends in the number of posts identified to more complex content analyses, most studies focused on sentiment analyses. Identifying sentiments toward vaccination expressed in social media can be useful to detect changes in beliefs and possible drops in confidence. Manual and automated methods of analyzing sentiments both come with their own benefits and limitations: manual coding may be easier and requires less technical skills than automated coding, whereas it is more prone to subjectivity biases and is time-consuming, and therefore does not allow for the analysis of a large number of posts. The complexity of coding discourse, particularly those using sarcasm or irony, is apparent with both automated and manual systems [40]. Researchers should choose analytical methods based on their personal objectives and resources available. Despite the possibility of using a combination of manual and automated coding or selecting automated systems that require less technical skills, more accurate and easy-to-use automated systems should also be developed.

The development of sentiment analysis as a tool in social media monitoring raises other challenges. Most studies identified in this review used simple binary categorizations of sentiments (eg, "negative" vs "positive"). However, discussions around vaccination tend to elicit complex sentiments, closely linked to deeper, more contextual themes of trust, confidence, and risk perception [131]. Categorizing sentiments as either negative or positive therefore fails to recognize nuances that would be crucial for the development of targeted responses to rebuild trust in vaccination. If automated coding systems are to be further developed, they need to take into account the nuances in sentiments around vaccination and move beyond the use of binary variables. More complex sentiment analysis will also improve the quality of the coding of videos, images, and emojis [132-135].

## Considerations for Future Research: Changing Digital Ecosystem

Following the Cambridge Analytica data misuse scandal and an increase in the amount of antivaccine content, Facebook announced a number of API changes aimed at better protecting user information between 2017 and 2019. These restrictions, along with GDPR laws, will pose restrictions on the type of data and research that can be performed on social media platforms and will require researchers to continuously adapt their methodologies [136,137]. Furthermore, platforms such as Pinterest, Facebook, and YouTube are responding to requests from public bodies to alter their content to respond to concerns about the spread of misinformation about vaccines and the presence of antivaccination content on social media [138-140]. These actions from social media platforms may change what users see but also what researchers study. Indeed, it may be that antivaccine groups move away from platforms that no longer monetize or make it easy for them to share information. Those with antivaccination views have not only been found to be using a mixture of websites and social media but also to migrate over to the dark web, where they are able to create and construct content-specific platforms from which their chosen ideologies can be shared [141]. The question of who should decide what content falls under antivaccination sentiment is also important. Social media platforms should work closely with vaccination experts to identify which posts to remove or keep, especially to avoid infringing on the public's freedom of expression.

## Study Limitations

Some limitations of this systematic scoping review should be acknowledged. Although articles in languages other than English were included for analysis, the search itself only used English keywords, which could have limited the results. Furthermore, the search strategy was comprehensive but did not include certain relevant keywords such as "infodemiology" or "infoveillance," which should be considered in future research. Data extraction was performed by three researchers, which could have caused inconsistencies even though the same data extraction sheet was used. Finally, social media monitoring constitutes a relatively new research field, which means that many real-life, practical uses of monitoring may have been omitted as they may not have been published in publicly available peer-reviewed journals or reports. It is also important

to note that as this is a fast-moving field, a high number of articles have been published since this review was performed, particularly around the COVID-19 pandemic. Methodologies for social media monitoring are expected to continuously and rapidly evolve, and further research should be performed to regularly update this review.

## Conclusion

Social media has changed the communication landscape around vaccination. The increasing use of social media by individuals to find and share information about vaccination, together with the growing volume of negative information about vaccination online, has influenced the way people assess the risks and benefits of vaccination. Social media monitoring studies have been developed with the aim of better understanding the type of information social media users are exposed to, and how this information is spread and shared across the world. This review has identified clear steps to perform social media monitoring that can be organized in three phases: (1) Preparation (defining the purpose of media monitoring, addressing ethical considerations); (2) Data extraction (selecting data extraction tools, developing comprehensive search strategies); and (3) Data analysis (geolocation, trends, content, sentiments, and reach). A wide range of tools for each of these steps have been identified in the literature but have not yet been evaluated. Therefore, to establish social media monitoring as a valuable research design, future research should aim to identify which methods are more robust and precise to extract and analyze data from social media.

## Conflicts of Interest

The LSHTM research group "The Vaccine Confidence Project" has received funding for other projects from the Bill & Melinda Gates Foundation, the Center for Strategic and International Studies, Innovative Medicines Initiative (IMI), National Institute for Health Research (UK), the World Health Organization, GlaxoSmithKline (GSK), Novartis, Johnson & Johnson, and Merck. HJL has done consulting on vaccine confidence with GSK and served on the Merck Vaccine Strategic Advisory Board. EK has received support for participating in meetings and advisory roundtables organised by GSK and Merck. The other authors have no conflicts of interest to declare.

## Multimedia Appendix 1

Search strategy developed for the EMBASE database.
[PDF File (Adobe PDF File), 522 KB-Multimedia Appendix 1]

## Multimedia Appendix 2

Summary of included articles.
[PDF File (Adobe PDF File), 622 KB-Multimedia Appendix 2]

## References

1. Poland GA, Jacobson RM. The age-old struggle against the antivaccinationists. N Engl J Med 2011 Jan 13;364(2):97-99. [doi: 10.1056/NEJMp1010594] [Medline: 21226573]
2. Stahl JP, Cohen R, Denis F, Gaudelus J, Martinot A, Lery T, et al. The impact of the web and social networks on vaccination. New challenges and opportunities offered to fight against vaccine hesitancy. Med Mal Infect 2016 May;46(3):117-122. [doi: 10.1016/j.medmal.2016.02.002] [Medline: 26987960]
3. Larson HJ, de Figueiredo A, Xiahong Z, Schulz WS, Verger P, Johnston IG, et al. The State of Vaccine Confidence 2016: Global Insights Through a 67-Country Survey. EBioMedicine 2016 Oct;12:295-301 [FREE Full text] [doi: 10.1016/j.ebiom.2016.08.042] [Medline: 27658738]
4. Larson HJ, Bertozzi S, Piot P. Redesigning the AIDS response for long-term impact. Bull World Health Organ 2011 Nov 01;89(11):846-852 [FREE Full text] [doi: 10.2471/BLT.11.087114] [Medline: 22084531]
5. Yaqub O, Castle-Clarke S, Sevdalis N, Chataway J. Attitudes to vaccination: a critical review. Soc Sci Med 2014 Jul;112:1-11 [FREE Full text] [doi: 10.1016/j.socscimed.2014.04.018] [Medline: 24788111]
6. Black S, Rappuoli R. A crisis of public confidence in vaccines. Sci Transl Med 2010 Dec 08;2(61):61mr1. [doi: 10.1126/scitranslmed.3001738] [Medline: 21148125]
7. Li HO, Bailey A, Huynh D, Chan J. YouTube as a source of information on COVID-19: a pandemic of misinformation? BMJ Glob Health 2020 May;5(5) [FREE Full text] [doi: 10.1136/bmjgh-2020-002604] [Medline: 32409327]
8. Blankenship EB, Goff ME, Yin J, Tse ZTH, Fu K, Liang H, et al. Sentiment, contents, and retweets: a study of two vaccine-related Twitter datasets. Perm J 2018;22:17-138 [FREE Full text] [doi: 10.7812/TPP/17-138] [Medline: 29911966]

XSL•FO
RenderX

9. Basch CH, MacLean SA. A content analysis of HPV related posts on instagram. Hum Vaccin Immunother 2019;15(7-8):1476-1478 [FREE Full text] [doi: 10.1080/21645515.2018.1560774] [Medline: 30570379]

10. Puri N, Coomes EA, Haghbayan H, Gunaratne K. Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. Hum Vaccin Immunother 2020 Nov 01;16(11):2586-2593. [doi: 10.1080/21645515.2020.1780846] [Medline: 32693678]

11. Konijn EA, Veldhuis J, Plaisier XS. YouTube as a research tool: three approaches. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):695-701. [doi: 10.1089/cyber.2012.0357] [Medline: 23659721]

12. Kelly J, Fealy G, Watson R. The image of you: constructing nursing identities in YouTube. J Adv Nurs 2012 Aug;68(8):1804-1813. [doi: 10.1111/j.1365-2648.2011.05872.x] [Medline: 22070735]

13. Pink S. Digital Ethnography: Principles and Practices. London, UK: SAGE Publications Ltd; 2016.

14. Brunson EK. The impact of social networks on parents' vaccination decisions. Pediatrics 2013 May;131(5):e1397-e1404. [doi: 10.1542/peds.2012-2452] [Medline: 23589813]

15. Johnson NF, Velásquez N, Restrepo NJ, Leahy R, Gabriel N, El Oud S, et al. The online competition between pro- and anti-vaccination views. Nature 2020 Jun;582(7811):230-233. [doi: 10.1038/s41586-020-2281-1] [Medline: 32499650]

16. Leask J, Kinnersley P, Jackson C, Cheater F, Bedford H, Rowles G. Communicating with parents about vaccination: a framework for health professionals. BMC Pediatr 2012 Sep 21;12:154 [FREE Full text] [doi: 10.1186/1471-2431-12-154] [Medline: 22998654]

17. Bahri P, Fogd J, Morales D, Kurz X, ADVANCE consortium. Application of real-time global media monitoring and 'derived questions' for enhancing communication by regulatory bodies: the case of human papillomavirus vaccines. BMC Med 2017 May 02;15(1):91 [FREE Full text] [doi: 10.1186/s12916-017-0850-4] [Medline: 28464900]

18. Steffens MS, Dunn AG, Wiley KE, Leask J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. BMC Public Health 2019 Oct 23;19(1):1348 [FREE Full text] [doi: 10.1186/s12889-019-7659-3] [Medline: 31640660]

19. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

20. Social media strategy development: a guide to using social media for public health communication. European Centre for Disease Prevention Control. Stockholm, Sweden: ECDC; 2016. URL: https://www.ecdc.europa.eu/en/publications-data/social-media-strategy-development-guide-using-social-media-public-health [accessed 2021-01-25]

21. Systematic scoping review on social media monitoring methods and interventions relating to vaccine hesitancy. European Centre for Disease Prevention and Control. Stockholm URL: https://www.ecdc.europa.eu/en/publications-data/systematic-scoping-review-social-media-monitoring-methods-and-interventions [accessed 2021-01-28]

22. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]

23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Method 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

24. Peters M, Godfrey C, McInerney P, Soares C, Khalil H, Parker D. The Joanna Briggs Institute Reviewers' Manual 2015: Methodology for JBI Scoping Reviews. Adelaide, SA Australia: The Joanna Briggs Institute; 2015. URL: https://nursing.lsuhsc.edu/JBI/docs/ReviewersManuals/Scoping-.pdf [accessed 2021-01-25]

25. Kaplan A, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Bus Horiz 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]

26. Skea Z. Communication about the MMR vaccine: an analysis of newspaper coverage and an internet discussion forum.: PhD thesis, University of Aberdeen; 2006. URL: https://eu03.alma.exlibrisgroup.com/view/delivery/44ABE_INST/12153284070005941 [accessed 2021-01-25]

27. Addawood A. Usage of scientific references in MMR vaccination debates on Twitter. 2018 Presented at: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); 2018; Barcelona p. 971-979. [doi: 10.1109/asonam.2018.8508385]

28. Bahk CY, Cumming M, Paushter L, Madoff LC, Thomson A, Brownstein JS. Publicly Available Online Tool Facilitates Real-Time Monitoring Of Vaccine Conversations And Sentiments. Health Aff (Millwood) 2016 Feb;35(2):341-347. [doi: 10.1377/hlthaff.2015.1092] [Medline: 26858390]

29. Becker BFH, Larson HJ, Bonhoeffer J, van Mulligen EM, Kors JA, Sturkenboom MCJM. Evaluation of a multinational, multilingual vaccine debate on Twitter. Vaccine 2016 Dec 07;34(50):6166-6171. [doi: 10.1016/j.vaccine.2016.11.007] [Medline: 27840012]

30. Bello-Orgaz G, Hernandez-Castro J, Camacho D. Detecting discussion communities on vaccination in twitter. Future Gener Comput Syst 2017 Jan;66:125-136. [doi: 10.1016/j.future.2016.06.032]

31. Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, et al. Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. Am J Public Health 2018 Oct;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]

32.    Chakraborty P, Colditz JB, Silvestre AJ, Friedman MR, Bogen KW, Primack BA, et al. Observation of public sentiment toward human papillomavirus vaccination on Twitter. Cogent Medicine 2017 Dec 12;4(1):1390853. [doi: 10.1080/2331205x.2017.1390853]

33.    Chen T, Dredze M. Vaccine images on Twitter: analysis of what images are shared. J Med Internet Res 2018 Apr 03;20(4):e130 [FREE Full text] [doi: 10.2196/jmir.8221] [Medline: 29615386]

34.    D'Andrea E, Ducange P, Bechini A, Renda A, Marcelloni F. Monitoring the public opinion about the vaccination topic from tweets analysis. Expert Syst Appl 2019 Feb;116:209-226. [doi: 10.1016/j.eswa.2018.09.009]

35.    Dredze M, Wood-Doughty Z, Quinn SC, Broniatowski DA. Vaccine opponents' use of Twitter during the 2016 US presidential election: Implications for practice and policy. Vaccine 2017 Aug 24;35(36):4670-4672. [doi: 10.1016/j.vaccine.2017.06.066] [Medline: 28688783]

36.    Du J, Tang L, Xiang Y, Zhi D, Xu J, Song H, et al. Public Perception Analysis of Tweets During the 2015 Measles Outbreak: Comparative Study Using Convolutional Neural Network Models. J Med Internet Res 2018 Jul 09;20(7):e236 [FREE Full text] [doi: 10.2196/jmir.9413] [Medline: 29986843]

37.    Du J, Xu J, Song H, Liu X, Tao C. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. J Biomed Semantics 2017 Mar 03;8(1):9 [FREE Full text] [doi: 10.1186/s13326-017-0120-6] [Medline: 28253919]

38.    Du J, Xu J, Song H, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):69 [FREE Full text] [doi: 10.1186/s12911-017-0469-6] [Medline: 28699569]

39.    Dunn AG, Leask J, Zhou X, Mandl KD, Coiera E. Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study. J Med Internet Res 2015 Jun 10;17(6):e144 [FREE Full text] [doi: 10.2196/jmir.4343] [Medline: 26063290]

40.    Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. Vaccine 2017 May 25;35(23):3033-3040 [FREE Full text] [doi: 10.1016/j.vaccine.2017.04.060] [Medline: 28461067]

41.    Huang X, Smith MC, Paul MJ, Ryzhkov D, Quinn SC, Broniatowski DA. Examining patterns of influenza vaccination in social media. In: AAAI Publications. 2017 Presented at: Thirty-First AAAI Conference on Artificial Intelligence; 2017; San Francisco.

42.    Kang GJ, Ewing-Nelson SR, Mackey L, Schlitt JT, Marathe A, Abbas KM, et al. Semantic network analysis of vaccine sentiment in online social media. Vaccine 2017 Jun 22;35(29):3621-3638 [FREE Full text] [doi: 10.1016/j.vaccine.2017.05.052] [Medline: 28554500]

43.    Kaptein R, Boertjes E, Langley D. Analyzing discussions on Twitter: Case study on HPV vaccinations. In: de Rijke M, Kenter T, de Vries AP, Zhai C, de Jong F, Radinsky K, et al, editors. Advances in Information Retrieval: 36th European Conference on IR Research, ECIR 2014, Amsterdam, The Netherlands, April 13-16, 2014. Proceedings. Switzerland: Springer, Cham; 2014:474-480.

44.    Keim-Malpass J, Mitchell EM, Sun E, Kennedy C. Using Twitter to Understand Public Perceptions Regarding the #HPV Vaccine: Opportunities for Public Health Nurses to Engage in Social Marketing. Public Health Nurs 2017 Jul;34(4):316-323. [doi: 10.1111/phn.12318] [Medline: 28261846]

45.    Krittanawong C, Tunhasiriwet A, Chirapongsathorn S, Kitai T. Tweeting influenza vaccine to cardiovascular health community. Eur J Cardiovasc Nurs 2017 Dec;16(8):704-706. [doi: 10.1177/1474515117707867] [Medline: 28440668]

46.    Love B, Himelboim I, Holton A, Stewart K. Twitter as a source of vaccination information: content drivers and what they are saying. Am J Infect Control 2013 Jun;41(6):568-570. [doi: 10.1016/j.ajic.2012.10.016] [Medline: 23726548]

47.    Mahoney LM, Tang T, Ji K, Ulrich-Schad J. The digital distribution of public health news surrounding the human papillomavirus vaccination: a longitudinal infodemiology study. JMIR Public Health Surveill 2015;1(1):e2 [FREE Full text] [doi: 10.2196/publichealth.3310] [Medline: 27227125]

48.    Massey PM, Budenz A, Leader A, Fisher K, Klassen AC, Yom-Tov E. What Drives Health Professionals to Tweet About #HPVvaccine? Identifying Strategies for Effective Communication. Prev Chronic Dis 2018 Feb 22;15:E26 [FREE Full text] [doi: 10.5888/pcd15.170320] [Medline: 29470166]

49.    Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC. Applying multiple data collection tools to quantify human papillomavirus vaccine communication on Twitter. J Med Internet Res 2016 Dec 05;18(12):e318 [FREE Full text] [doi: 10.2196/jmir.6670] [Medline: 27919863]

50.    McNeill A, Harris PR, Briggs P. Twitter Influence on UK Vaccination and Antiviral Uptake during the 2009 H1N1 Pandemic. Front Public Health 2016;4:26. [doi: 10.3389/fpubh.2016.00026] [Medline: 26942174]

51.    Mitra T, Counts S, Pennebaker JW. Understanding anti-vaccination attitudes in social media. In: Association for the Advancement of Artificial Intelligence. 2016 Presented at: Tenth International AAAI Conference on Web and Social Media; 2016; Cologne, Germany.

52.    Numnark S, Ingsriswang S, Wichadakul D. VaccineWatch: A monitoring system of vaccine messages from social media data. 2014 Presented at: The 8th International Conference on Systems Biology; October 24-27, 2014; Qingdao, China p. 112-117. [doi: 10.1109/isb.2014.6990742]

53. Porat T, Garaizar P, Ferrero M, Jones H, Ashworth M, Vadillo MA. Content and source analysis of popular tweets following a recent case of diphtheria in Spain. Eur J Public Health 2019 Feb 01;29(1):117-122. [doi: 10.1093/eurpub/cky144] [Medline: 30084926]

54. Radzikowski J, Stefanidis A, Jacobsen KH, Croitoru A, Crooks A, Delamater PL. The measles vaccination narrative in Twitter: a quantitative analysis. JMIR Public Health Surveill 2016;2(1):e1 [FREE Full text] [doi: 10.2196/publichealth.5059] [Medline: 27227144]

55. Sanawi J, Samani M, Taibi M. #Vaccination: Identifying influencers in the vaccination discussion on twitter through social network visualisation. Int J Bus Soc 2017;18(S4):718-726.

56. Shapiro GK, Surian D, Dunn AG, Perry R, Kelaher M. Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. BMJ Open 2017 Oct 05;7(10):e016869 [FREE Full text] [doi: 10.1136/bmjopen-2017-016869] [Medline: 28982821]

57. Smith MC, Dredze M, Quinn SC, Broniatowski DA. Monitoring real-time spatial public health discussions in the context of vaccine hesitancy. 2017 Presented at: SMM4H@AMIA; 2017; Washington, DC.

58. Surian D, Nguyen DQ, Kennedy G, Johnson M, Coiera E, Dunn AG. Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. J Med Internet Res 2016 Aug 29;18(8):e232 [FREE Full text] [doi: 10.2196/jmir.6045] [Medline: 27573910]

59. Tang L, Bie B, Zhi D. Tweeting about measles during stages of an outbreak: A semantic network approach to the framing of an emerging infectious disease. Am J Infect Control 2018 Dec;46(12):1375-1380 [FREE Full text] [doi: 10.1016/j.ajic.2018.05.019] [Medline: 29929837]

60. Teoh D, Shaikh R, Vogel RI, Zoellner T, Carson L, Kulasingam S, et al. A Cross-Sectional Review of Cervical Cancer Messages on Twitter During Cervical Cancer Awareness Month. J Low Genit Tract Dis 2018 Jan;22(1):8-12 [FREE Full text] [doi: 10.1097/LGT.0000000000000363] [Medline: 29271850]

61. Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. Soc Sci Med 2017 Oct;191:168-175 [FREE Full text] [doi: 10.1016/j.socscimed.2017.08.041] [Medline: 28926775]

62. Yuan X, Crooks A. Examining online vaccination discussion and communities in Twitter. New York: Association for Computing Machinery; 2018 Presented at: SMSociety '18: International Conference on Social Media and Society; July 18-20, 2018; Copenhagen, Denmark p. 197-206. [doi: 10.1145/3217804.3217912]

63. D'Andrea E, Ducange P, Marcelloni F. Monitoring negative opinion about vaccines from tweets analysis. : IEEE; 2017 Presented at: Third international conference on research in computational intelligence and communication networks (ICRCICN); 2017; Kolkata p. 186-191. [doi: 10.1109/icrcicn.2017.8234504]

64. Ninkov A, Vaughan L. A webometric analysis of the online vaccination debate. J Assoc Inf Sci Technol 2017 Mar 20;68(5):1285-1294. [doi: 10.1002/asi.23758]

65. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol 2011 Oct;7(10):e1002199 [FREE Full text] [doi: 10.1371/journal.pcbi.1002199] [Medline: 22022249]

66. Pananos AD, Bury TM, Wang C, Schonfeld J, Mohanty SP, Nyhan B, et al. Critical dynamics in population vaccinating behavior. Proc Natl Acad Sci U S A 2017 Dec 26;114(52):13762-13767 [FREE Full text] [doi: 10.1073/pnas.1704093114] [Medline: 29229821]

67. Zhou X, Coiera E, Tsafnat G, Arachi D, Ong M, Dunn A. Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter. Stud Health Technol Inform 2015;216:761-765. [Medline: 26262154]

68. Ache KA, Wallace LS. Human papillomavirus vaccination coverage on YouTube. Am J Prev Med 2008 Oct;35(4):389-392. [doi: 10.1016/j.amepre.2008.06.029] [Medline: 18675530]

69. Basch CH, Hillyer GC, Berdnik A, Basch CE. YouTube™ videos related to human papillomavirus: the need for professional communication. Int J Adolesc Med Health 2016 Apr 09;30(1). [doi: 10.1515/ijamh-2015-0122] [Medline: 27060736]

70. Basch CH, Zybert P, Reeves R, Basch CE. What do popular YouTube videos say about vaccines? Child Care Health Dev 2017 Jul;43(4):499-503. [doi: 10.1111/cch.12442] [Medline: 28105642]

71. Briones R, Nan X, Madden K, Waks L. When vaccines go viral: an analysis of HPV vaccine coverage on YouTube. Health Commun 2012;27(5):478-485. [doi: 10.1080/10410236.2011.610258] [Medline: 22029723]

72. Cambra U, Díaz V, Herrero S. Vacunas y anti vacunas en la red social Youtube. Opción: Revista de Ciencias Humanas y Sociales 2016;32:447-465.

73. Covolo L, Ceretti E, Passeri C, Boletti M, Gelatti U. What arguments on vaccinations run through YouTube videos in Italy? A content analysis. Hum Vaccin Immunother 2017 Jul 03;13(7):1693-1699 [FREE Full text] [doi: 10.1080/21645515.2017.1306159] [Medline: 28362544]

74. Donzelli G, Palomba G, Federigi I, Aquino F, Cioni L, Verani M, et al. Misinformation on vaccination: A quantitative analysis of YouTube videos. Hum Vaccin Immunother 2018 Jul 03;14(7):1654-1659 [FREE Full text] [doi: 10.1080/21645515.2018.1454572] [Medline: 29553872]

75.  Ekram S, Debiec KE, Pumper MA, Moreno MA. Content and Commentary: HPV Vaccine and YouTube. J Pediatr Adolesc Gynecol 2019 Apr;32(2):153-157. [doi: 10.1016/j.jpag.2018.11.001] [Medline: 30445163]

76.  Hernández-García I, Fernández Porcel C. Characteristics of YouTube™ videos in Spanish about the vaccine against meningococcus B. Vacunas (English Edition) 2018 Jul;19(2):37-43. [doi: 10.1016/j.vacune.2018.11.002]

77.  Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. JAMA 2007 Dec 05;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]

78.  Tuells J, Martínez-Martínez PJ, Duro-Torrijos JL, Caballero P, Fraga-Freijeiro P, Navarro-López V. [Characteristics of the Videos in Spanish Posted on Youtube about Human Papillomavirus Vaccines]. Rev Esp Salud Publica 2015;89(1):107-115 [FREE Full text] [doi: 10.4321/S1135-57272015000100012] [Medline: 25946591]

79.  Venkatraman A, Garg N, Kumar N. Greater freedom of speech on Web 2.0 correlates with dominance of views linking vaccines to autism. Vaccine 2015 Mar 17;33(12):1422-1425. [doi: 10.1016/j.vaccine.2015.01.078] [Medline: 25665960]

80.  Buchanan R, Beckett RD. Assessment of vaccination-related information for consumers available on Facebook. Health Info Libr J 2014 Sep;31(3):227-234. [doi: 10.1111/hir.12073] [Medline: 25041499]

81.  Faasse K, Chatman CJ, Martin LR. A comparison of language use in pro- and anti-vaccination comments in response to a high profile Facebook post. Vaccine 2016 Nov 11;34(47):5808-5814. [doi: 10.1016/j.vaccine.2016.09.029] [Medline: 27707558]

82.  Furini M, Menegoni G. Public Health and Social Media: Language Analysis of Vaccine Conversations. New York: IEEE; 2018 Presented at: International workshop on social sensing (SocialSens); Apr 17, 2018; Orlando, FL p. 50-55.

83.  Luisi MLR. Who Gives a "Like" About the HPV Vaccine? Kansan Parent/Guardian Perceptions and Social Media Representations. Doctoral dissertation, University of Kansas. 2017. URL: https://kuscholarworks.ku.edu/bitstream/handle/1808/26037/Luisi_ku_0099D_15291_DATA_1.pdf?sequence=1 [accessed 2021-01-25]

84.  Ma J, Stahl L. A multimodal critical discourse analysis of anti-vaccination information on Facebook. Libr Inf Sci Res 2017 Oct;39(4):303-310. [doi: 10.1016/j.lisr.2017.11.005]

85.  Orr D, Baram-Tsabari A. Science and Politics in the Polio Vaccination Debate on Facebook: A Mixed-Methods Approach to Public Engagement in a Science-Based Dialogue. J Microbiol Biol Educ 2018;19(1) [FREE Full text] [doi: 10.1128/jmbe.v19i1.1500] [Medline: 29904532]

86.  Orr D, Baram-Tsabari A, Landsman K. Social media as a platform for health-related public debates and discussions: the Polio vaccine on Facebook. Isr J Health Policy Res 2016;5:34 [FREE Full text] [doi: 10.1186/s13584-016-0093-4] [Medline: 27843544]

87.  Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine 2018 Jun 14;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]

88.  Smith N, Graham T. Mapping the anti-vaccination movement on Facebook. Inf Commun Soc 2017 Dec 27;22(9):1310-1327. [doi: 10.1080/1369118x.2017.1418406]

89.  Suragh TA, Lamprianou S, MacDonald NE, Loharikar AR, Balakrishnan MR, Benes O, et al. Cluster anxiety-related adverse events following immunization (AEFI): An assessment of reports detected in social media and those identified using an online search engine. Vaccine 2018 Sep 25;36(40):5949-5954 [FREE Full text] [doi: 10.1016/j.vaccine.2018.08.064] [Medline: 30172632]

90.  Tustin JL, Crowcroft NS, Gesink D, Johnson I, Keelan J, Lachapelle B. User-driven comments on a Facebook advertisement recruiting Canadian parents in a study on immunization: content analysis. JMIR Public Health Surveill 2018 Sep 20;4(3):e10090 [FREE Full text] [doi: 10.2196/10090] [Medline: 30249585]

91.  Vayreda A, Antaki C. To vaccinate or not? The disqualification of commercial sources of health advice in an online forum. Commun Med 2011;8(3):273-282. [doi: 10.1558/cam.v8i3.273] [Medline: 23264989]

92.  Goh D, Chi J. Central or peripheral? Information elaboration cues on childhood vaccination in an online parenting forum. Comput Hum Behav 2017 Apr;69:181-188. [doi: 10.1016/j.chb.2016.11.066]

93.  Penţa MA, Băban A. Dangerous agent or saviour? HPV vaccine representations on online discussion forums in Romania. Int J Behav Med 2014 Feb;21(1):20-28. [doi: 10.1007/s12529-013-9340-z] [Medline: 24057409]

94.  Fadda M, Allam A, Schulz PJ. Arguments and sources on Italian online forums on childhood vaccinations: Results of a content analysis. Vaccine 2015 Dec 16;33(51):7152-7159. [doi: 10.1016/j.vaccine.2015.11.007] [Medline: 26592140]

95.  Nicholson MS, Leask J. Lessons from an online debate about measles-mumps-rubella (MMR) immunization. Vaccine 2012 May 28;30(25):3806-3812. [doi: 10.1016/j.vaccine.2011.10.072] [Medline: 22063388]

96.  Luoma‐aho V, Tirkkonen P, Vos M. Monitoring the issue arenas of the swine‐flu discussion. JCOM 2013 Jul 26;17(3):239-251. [doi: 10.1108/jcom-11-2010-0069]

97.  Tangherlini TR, Roychowdhury V, Glenn B, Crespi CM, Bandari R, Wadia A, et al. "Mommy Blogs" and the vaccination exemption narrative: results from a machine-learning approach for story aggregation on parenting social media sites. JMIR Public Health Surveill 2016 Nov 22;2(2):e166 [FREE Full text] [doi: 10.2196/publichealth.6586] [Medline: 27876690]

98.  Skeppstedt M, Kerren A, Stede M. Vaccine hesitancy in discussion forums: computer-assisted argument mining with topic models. Stud Health Technol Inform 2018;247:366-370. [Medline: 29677984]

99.  Guidry JP, Carlyle K, Messner M, Jin Y. On pins and needles: how vaccines are portrayed on Pinterest. Vaccine 2015 Sep 22;33(39):5051-5056. [doi: 10.1016/j.vaccine.2015.08.064] [Medline: 26319742]

100.  Chen B, Zhang JM, Jiang Z, Shao J, Jiang T, Wang Z, et al. Media and public reactions toward vaccination during the 'hepatitis B vaccine crisis' in China. Vaccine 2015 Apr 08;33(15):1780-1785. [doi: 10.1016/j.vaccine.2015.02.046] [Medline: 25731787]

101.  Rivera I, Warren J, Curran J. Quantifying mood, content and dynamics of health forums. In: ACSW '16: Proceedings of the Australasian Computer Science Week Multiconference. New York, USA: Association for Computing Machinery; 2016 Presented at: Australasian Computer Science Week Multiconference; February 2016; Canberra, Australia. [doi: 10.1145/2843043.2843379]

102.  Lopez C, Krauskopf E, Villota C, Burzio L, Villegas J. Cervical cancer, human papillomavirus and vaccines: assessment of the information retrieved from general knowledge websites in Chile. Public Health 2017 Jul;148:19-24. [doi: 10.1016/j.puhe.2017.02.017] [Medline: 28404529]

103.  Nawa N, Kogaki S, Takahashi K, Ishida H, Baden H, Katsuragi S, et al. Analysis of public concerns about influenza vaccinations by mining a massive online question dataset in Japan. Vaccine 2016 Jun 08;34(27):3207-3213. [doi: 10.1016/j.vaccine.2016.01.008] [Medline: 26776467]

104.  Seeman N, Ing A, Rizo C. Assessing and responding in real time to online anti-vaccine sentiment during a flu pandemic. Healthc Q 2010;13 Spec No:8-15. [doi: 10.12927/hcq.2010.21923] [Medline: 20959725]

105.  Deiner MS, Fathy C, Kim J, Niemeyer K, Ramirez D, Ackley SF, et al. Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics J 2019 Sep;25(3):1116-1132 [FREE Full text] [doi: 10.1177/1460458217740723] [Medline: 29148313]

106.  Sundstrom B, Aylor E, Cartmell KB, Brandt HM, Bryant DC, Hughes Halbert C, et al. Beyond the birds and the bees: a qualitative content analysis of online HPV vaccination communication. J Commun Healthc 2018 Jun 06;11(3):205-214. [doi: 10.1080/17538068.2018.1484984]

107.  Mollema L, Harmsen IA, Broekhuizen E, Clijnk R, De Melker H, Paulussen T, et al. Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013. J Med Internet Res 2015 May 26;17(5):e128 [FREE Full text] [doi: 10.2196/jmir.3863] [Medline: 26013683]

108.  Jang SM, Mckeever BW, Mckeever R, Kim JK. From Social Media to Mainstream News: The Information Flow of the Vaccine-Autism Controversy in the US, Canada, and the UK. Health Commun 2019 Jan;34(1):110-117. [doi: 10.1080/10410236.2017.1384433] [Medline: 29028371]

109.  Aquino F, Donzelli G, De Franco E, Privitera G, Lopalco PL, Carducci A. The web and public confidence in MMR vaccination in Italy. Vaccine 2017 Aug 16;35(35 Pt B):4494-4498. [doi: 10.1016/j.vaccine.2017.07.029] [Medline: 28736200]

110.  Lehmann BA, Ruiter RA, Kok G. A qualitative study of the coverage of influenza vaccination on Dutch news sites and social media websites. BMC Public Health 2013 Jun 05;13:547 [FREE Full text] [doi: 10.1186/1471-2458-13-547] [Medline: 23738769]

111.  Zheleva E, Getoor L. To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In: Proceedings of the 18th international conference on world wide web. 2009 Presented at: 18th international conference on World wide web; April 20-24, 2009; Madrid, Spain. [doi: 10.1145/1526709.1526781]

112.  Lea D, Bradbery J. Oxford Advanced Learner's Dictionary. Oxford: Oxford University Press; 2020.

113.  Boyd D, Crawford K. Critical questions for big data. Inf Commun Soc 2012 Jun;15(5):662-679. [doi: 10.1080/1369118x.2012.678878]

114.  Lewandowsky S, Ecker U, Cook J. Beyond misinformation: understanding and coping with the "post-truth" era. J Appl Res Mem Cogn 2017 Dec;6(4):353-369. [doi: 10.1016/j.jarmac.2017.07.008]

115.  Dredze M, Broniatowski DA, Smith MC, Hilyard KM. Understanding vaccine refusal: why we need Social media now. Am J Prev Med 2016 Apr;50(4):550-552 [FREE Full text] [doi: 10.1016/j.amepre.2015.10.002] [Medline: 26655067]

116.  Cihon P, Yasseri T. A biased review of biases in Twitter studies on political collective action. Front Phys 2016 Aug 09;4(34). [doi: 10.3389/fphy.2016.00034]

117.  Kim Y, Huang J, Emery S. Garbage in, garbage out: data collection, quality assessment and reporting standards for social media data use in health research, infodemiology and digital disease detection. J Med Internet Res 2016 Feb 26;18(2):e41 [FREE Full text] [doi: 10.2196/jmir.4738] [Medline: 26920122]

118.  Hargittai E. Potential biases in big data: omitted voices on social media. Social Science Computer Review 2018 Jul 30;38(1):10-24. [doi: 10.1177/0894439318788322]

119.  Beninger K, Fry A, Jago N, Lepps H, Nass L, Silvester H. Research using social media; users' views. London: NatCen Social Research; 2014. URL: https://www.bl.uk/britishlibrary/~/media/bl/global/social-welfare/pdfs/non-secure/r/e/s/research-using-social-media-users-views.pdf [accessed 2021-01-25]

120.  Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):708-713 [FREE Full text] [doi: 10.1089/cyber.2012.0334] [Medline: 23679571]

121.  González F, Yu Y, Figueroa A, López C, Aragon C. Global Reactions to the Cambridge Analytica Scandal: An Inter-Language Social Media Study. 2019 Presented at: WWW '19: The Web Conference; May 2019; San Francisco. [doi: 10.1145/3308560.3316456]

122. GDPR – the fabric of a success story. European Commission. Brussels, Belgium; 2020. URL: https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules/gdpr-fabric-success-story_en [accessed 2021-01-25]

123. Carr I, Zimmermann K, van't Veer P, Teare H. Ethical design: At the Interface of Ethics for Big Data and the European Union's General Data Protection Regulation. European Commission. 2018. URL: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ba270ebd&appId=PPGMS [accessed 2021-01-25]

124. Driscoll K, Walker S. Big data, big questions. Working within a black box: Transparency in the collection and production of big twitter data. Int J Commun 2014;8:20 [FREE Full text]

125. Morstatter F, Pfeffer J, Liu H, Carley K. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. 2013 Presented at: Seventh international AAAI conference on weblogs and social media; July 8-11, 2013; Cambridge, MA.

126. Halfpenny P, Procter R. In: Seaman J, editor. Innovations in digital research method. London: Sage; 2015.

127. González-Bailón S, Borge-Holthoefer J, Rivero A, Moreno Y. The dynamics of protest recruitment through an online network. Sci Rep 2011;1:197. [doi: 10.1038/srep00197] [Medline: 22355712]

128. Choi TR, Sung Y. Instagram versus Snapchat: Self-expression and privacy concern on social media. Telemat Inform 2018 Dec;35(8):2289-2298. [doi: 10.1016/j.tele.2018.09.009]

129. Tufekci Z. Big questions for social media big data: Representativeness, validity and other methodological pitfalls. 2014 Presented at: Eighth International AAAI Conference on Weblogs and Social Media; June 1-4, 2014; Michigan, USA.

130. Gray V. Learning to look: Pulsar partners with Manchester School of Arts to understand the difference between how machines and humans look at images. Pulsar. 2018. URL: https://www.pulsarplatform.com/blog/2018/learning-to-look-pulsar-partners-with-manchester-school-of-arts-to-understand-the-difference-between-how-machines-and-humans-look-at-images/ [accessed 2021-01-25]

131. Gaspar R, Pedro C, Panagiotopoulos P, Seibt B. Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. Comput Hum Behav 2016 Mar;56:179-191. [doi: 10.1016/j.chb.2015.11.040]

132. Thelwall M. Social media analytics for YouTube comments: potential and limitations. Int J Soc Res Method 2017 Sep 21;21(3):303-316. [doi: 10.1080/13645579.2017.1381821]

133. Majumder N, Hazarika D, Gelbukh A, Cambria E, Poria S. Multimodal sentiment analysis using hierarchical fusion with context modeling. Knowl Based Syst 2018 Dec;161:124-133. [doi: 10.1016/j.knosys.2018.07.041]

134. Amarasekara I, Grant WJ. Exploring the YouTube science communication gender gap: A sentiment analysis. Public Underst Sci 2019 Jan;28(1):68-84. [doi: 10.1177/0963662518786654] [Medline: 29974815]

135. Lama Y, Chen T, Dredze M, Jamison A, Quinn SC, Broniatowski DA. Discordance between human papillomavirus Twitter images and disparities in human papillomavirus risk and disease in the United States: mixed-methods analysis. J Med Internet Res 2018 Sep 14;20(9):e10244 [FREE Full text] [doi: 10.2196/10244] [Medline: 30217792]

136. Wong JC. Facebook's Zuckerberg announces privacy overhaul: We don't have the strongest reputation. The Guardian. 2019 Apr 30. URL: https://www.theguardian.com/technology/2019/apr/30/facebook-f8-conference-privacy-mark-zuckerberg [accessed 2021-01-25]

137. Rieder B. Studying Facebook via data extraction: the Netvizz application. 2013 Presented at: 5th annual ACM web science conference; May 2013; Paris, France. [doi: 10.1145/2464464.2464475]

138. Wilson M. The tech giant fighting anti-vaxxers isn't Twitter or Facebook. It's Pinterest. Fast Company. 2019. URL: https://www.fastcompany.com/90310970/the-tech-giant-fighting-anti-vaxxers-isnt-twitter-or-facebook-its-pinterest [accessed 2021-01-25]

139. Wong JC. How Facebook and YouTube help spread anti-vaxxer propaganda. The Guardian. 2019 Feb 01. URL: https://www.theguardian.com/media/2019/feb/01/facebook-youtube-anti-vaccination-misinformation-social-media [accessed 2021-01-25]

140. YouTube takes ads off 'anti-vax' video channels. BBC. 2019 Feb 25. URL: https://www.bbc.com/news/technology-47357252 [accessed 2021-01-25]

141. Allen C. Extreme far right: 'pick'n'mix' ideologies and direct messaging online make for deadly new combination. The Conversation. 2019 Mar 19. URL: https://theconversation.com/extreme-far-right-picknmix-ideologies-and-direct-messaging-online-make-for-deadly-new-combination-112331 [accessed 2021-01-25]

## Abbreviations

**API:** application programming interface
**ECDC:** European Centre for Disease Prevention and Control
**GDPR:** General Data Protection Regulation
**HPV:** human papillomavirus
**VCP:** Vaccine Confidence Project

XSL•FO
**RenderX**