



OPEN ACCESS

EDITED BY

Rodrigo Muiño,
Veterinary Centre Meira, Spain

REVIEWED BY

Birgit Ursula Stetina,
Sigmund Freud University Vienna, Austria
David Andrew Singleton,
University of Liverpool, United Kingdom

*CORRESPONDENCE

Kirsten M. McMillan
✉ kirsten.mcmillan@dogstrust.org.uk

SPECIALTY SECTION

This article was submitted to
Animal Behavior and Welfare,
a section of the journal
Frontiers in Veterinary Science

RECEIVED 19 October 2022

ACCEPTED 06 February 2023

PUBLISHED 01 March 2023

CITATION

McMillan KM, Anderson KL and Christley RM
(2023) Pooches on a platform: Text mining
twitter for sector perceptions of dogs during a
global pandemic. *Front. Vet. Sci.* 10:1074542.
doi: 10.3389/fvets.2023.1074542

COPYRIGHT

© 2023 McMillan, Anderson and Christley. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Pooches on a platform: Text mining twitter for sector perceptions of dogs during a global pandemic

Kirsten M. McMillan*, Katharine L. Anderson and
Robert M. Christley

Dogs Trust, London, United Kingdom

Introduction: Businesses commonly text mine Twitter data to identify patterns and extract valuable information. However, this method is rarely applied to the animal welfare sector. Here, we describe Twitter conversations regarding dogs during a global pandemic, assess the evolution of sentiment, and examine the dynamics of sector influence.

Methods: Between March and August 2020, we gathered 61,088 unique tweets from the United Kingdom and Republic of Ireland, relating to COVID-19 and dogs. Tweets were assigned to one of four pandemic phases and active accounts were assigned to a sector: Personal (i.e., UK and ROI public), Press (i.e., mass media), State (i.e., Government, Police, and NHS), and Other (i.e., welfare organizations, social enterprises, research organizations, charity, and business).

Results: Word frequency and sentiment analysis between phases and sectors were assessed, and cross correlation functions and lagged regressions were used to evaluate sector influence. Topical foci of conversations included: meat trade, separation anxiety and dog theft. Sentiment score remained stable until the last phase where sentiment decreased ($F_{3,78,508} = 44.4, p < 0.001$), representing an increased use of negative language. Sentiment differed between the four sectors ($F_{3,11,794} = 52.2, p < 0.001$), with Personal and Press accounts presenting the greatest use of negative language. Personal accounts were initially partly influenced by State accounts ($R = -0.26; p = 0.05$), however this altered to Press accounts by the last phase ($R = -0.31; p = 0.02$).

Discussion: Our findings highlight that whilst Personal accounts may affect sector-specific messaging online, perhaps more importantly: language used, and sentiment expressed by Press, State and Other accounts may influence public perception. This draws attention to the importance of sector responsibility regarding accurate and appropriate messaging, as irresponsible/ill-considered comments or campaigns may impact future human-animal interaction.

KEYWORDS

dog, companion animal, COVID-19, Twitter, web scraped data, text mining, sentiment analysis

Introduction

With the rapid development of the internet and mobile networks, social media platforms have grown quickly (1), creating a vast source of raw data. Computational tools, such as text mining, allow researchers to collate and analyze user-generated language data, providing insight into online behavior and cohort opinion. Whilst charities may use these platforms to help promote their message and gain supporters, they rarely take advantage of computational tools to better understand their audience. As one of the most powerful

social media platforms, Twitter allows both consumers and brands to deliver insights and optimize engagement (2). Here, we argue that these tools may provide benefit to the animal welfare sector, as social media creates an environment where animals are represented within human society.

Twitter, established in 2006, is a microblogging social media platform that allows users to “tweet” content publicly, in 280 characters or less. With ~330 million active users, ranging from personal to government sanctioned accounts, Twitter is arguably the most utilized short-text discussion forum in the world (1). It can disseminate and reflect information broadly and rapidly, allowing for single or multi-way communication. Text mining (3), i.e., the ability to search for text or hashtags (keywords related to a topic that are preceded by #), enables researchers to examine Twitter archives and extract relevant information. This not only allows for informed decision-making by highlighting current areas of interest or concern, but it also provides situational awareness and flexibility, which may be of particular interest at present due to the unprecedented impacts of the novel coronavirus termed COVID-19 [caused by SARS-CoV-2 (4)].

Sentiment analysis is a text mining technique that uses machine learning and natural language processing to examine the content of free text for the intensity of positive and negative opinions and emotions. Computerized packages have been developed that automate the process, allowing large numbers of free-text comments to be quickly processed into quantitative sentiment scores. Lexicon-based sentiment analysis methods, e.g., *nrc*, *afinn*, and *bing*, have become very popular due to their unsupervised nature and easy-use properties (5). The immediate nature of Twitter feeds enables communication in real-time between peers; government and public sector; companies and suppliers; industry and consumers; charities and supporters; and many more, providing a wide-ranging source of user-generated language data. While businesses may use Twitter as a cost-effective method of engaging with their consumer base, personal accounts may present more emotive language, possibly supplemented by information sourced from press or government accounts. Consequently, language used, and sentiment presented may vary across sectors, which in turn, may have a critical impact on the behavior of others, particularly during times of crisis (6). Understanding the influence that sectors have on public perception, and *vice versa*, is imperative to animal welfare issues and those working within the field, as this will impact upon public engagement and future policy making. Whilst text mining and sentiment analysis of social media data has been widely employed within multiple fields, such as politics (7, 8), disaster monitoring and response (9, 10), ecological modeling (11, 12), and human health (13, 14), this tool remains relatively underutilized within animal welfare (15, 16).

As a result of the COVID-19 pandemic, countries worldwide have faced difficulties ranging from human health services (17) to animal conservation (18). Three main impacts have been suggested regarding animal welfare (19): (1) an immediate impact due to sudden human confinement and inactivity; (2) a medium to long-term impact due to the effects of the resulting economic crisis on farming and veterinary services; and (3) an increased attention to the public health implications of coronavirus infections in animals, including those found wild (20), in farms (21), and

those living with humans, i.e., pets (22). With regards to pet dogs, there is growing concern regarding the impact that the pandemic has had on multiple aspects of their welfare (23–27). These include (but are not limited to): deviations from the daily routine, increasing the probability of a pet dog developing behavioral issues; veterinary visits falling due to real or perceived difficulty in accessing veterinary care; and an increased demand for puppies and associated intensification of large-scale breeding and/or rise of illegal puppy smuggling. Tracking main topics discussed online and assessing sentiment surrounding these topics may be beneficial to multiple stakeholders as it allows for the development of proactive response strategies (28).

To ensure progress within the field, it is imperative that methods of assessing and measuring topics of concern are continually identified and developed. As far as the authors are aware, no study has attempted to examine sentiment expressed by specific sectors on social media, within the context of animal welfare. Given the low-cost of data utilization, ability to search archives for relevant topics and real-time streaming nature of Twitter, these methods may provide a unique, yet untapped, tool for the animal welfare sector. Consequently, we focused on canine welfare as an example topic, to assess the functionality of applying text mining and sentiment analyses to tweets, in order to gain meaningful temporal and sector specific insight. More specifically, we aimed to: (1) highlight the topical foci of conversations regarding canine-COVID topics in the United Kingdom and Republic of Ireland, (2) assess temporal evolution of sentiment (throughout key phases of the pandemic), and (3) assess and compare sentiment expressed by specific sectors (“Personal”, “Press”, “State”, and “Other”) and examine their influence on each other.

Materials and methods

Data collection

Implementing R package *rtweet* (29), we connected with the official Twitter application programming interface (API) and gathered English language tweets from the United Kingdom (UK) and the Republic of Ireland (ROI) by applying a geocode buffer (Lat = 55.5166, Long = -4.0661, Buffer = 390 miles; [Supplementary Figure 1](#)). Tweets were searched using two term lists: (1) words related to COVID-19 and (2) words related to dogs. Tweets where ≥ 1 search term(s) from both lists were present, were included in the analysis. All search terms and their reported frequencies are listed in [Supplementary Table 1](#). The search engine relies on Twitter’s search API and is limited to tweets published in the past 7 days. Consequently, the script was run weekly. We began running the script on the March 27th, 2020, purging retweets and storing backup datafiles incrementally, in order to preserve historical data (further details on dates included below).

Data cleaning and manipulation

Data analysis was performed with the statistical software R (30) (version 4.0.2), using the aforementioned package

rtweet (29), as well as *dplyr* (31), *stringr* (32), and *tidyverse* (33) for data cleaning and manipulation. Packages *tidytext* (34) and *tm* (35) were used for text mining. For the production of figures, *ggplot2* (36) was implemented, along with *wordcloud* (37).

A profane word list was imported (38) and augmented by the authors, to include 750 English terms that could be found offensive. Profanities were then removed from tweets, as sentiment of profane language can be positive or negative, depending on context. For example, profanity may co-occur with insults or abusive speech, or may be used in tweets to express emphasis regarding positive sentiment. Duplicate tweets were identified and removed, by matching Twitter handle (i.e., account name), content of tweet and date. To ensure that the tweet was relevant, i.e., referred to the COVID-19 and dogs, relevance was checked by searching for the presence of one or more pre-defined terms (Supplementary Table 2). Tweets that did not include any relevant words were excluded from the analysis.

The UK entered lockdown on March 23rd, 2020 (39). Data collection for this research commenced on the March 27th, 2020, succeeding reactive project initiation and planning. Data collection ended on the August 22nd, 2020: the date where the daily number of relevant tweets first fell below 100 in total. This period will now be referred to as the “full period”. Within the full period, tweets were assigned to one of four phases: (1) Lockdown (LD; March 27th–May 12th, 2020), (2) Phase Ease 1 (PE1; May 13th–May 31st, 2020), (3) Phase Ease 2 (PE2; June 1st–July 3rd, 2020), and (4) Phase Ease 3 (PE3; July 4th–August 22nd, 2020). Following LD, the devolved nations (Northern Ireland, Wales, and Scotland) and ROI followed their own paths with regards to determining restrictions and, in turn, easing restrictions: based on their respective number of cases and regional NHS capacity (40). Unfortunately, whilst data were reliably collected from the UK and ROI, tweet location could not be assigned to country as most Twitter users opt to keep their location unpublished (41). Consequently, for the purposes of this study, we broadly determined four pandemic phases based on major events occurring across the UK and ROI (39, 40, 42–57), details for which are outlined in Figure 1.

Twitter accounts where 4 or more relevant tweets were posted during the full period, were assigned to one of four sectors: (1) ‘Personal’, i.e., personal accounts/UK and ROI public; (2) ‘Press’, i.e., all variations of mass media; (3) ‘State’, i.e., Government, Police, and NHS; and (4) ‘Other’, i.e., all other sectors such as (but not limited to) animal welfare organizations, social enterprises, research organizations, charity and business. Sector categorization was carried out *via* manual assessment of the Twitter account “Bio”: a small public summary regarding account holder or business. The probability of these accounts functioning as bots or bot accounts, i.e., automated programs, were assessed using R package *tweetbotornot* (58). Legitimate bots may generate a large number of benign tweets delivering news and updating feeds, while malicious bots may spread spam, incorrect and/or irresponsible content (59). However, this was for reference only, and no data were excluded due to the outcome. Limitations of the above methodologies are discussed in Supplementary Note 1.

Word frequency and negation

Word frequency was assessed and compared using the full corpus of tweets (tweet content and hashtag/s), where punctuation, stop words (i.e., commonly used words which do not add much meaning to a sentence, e.g., “the”, “an”, “in”, “just”, “can” etc.) and numbers were removed. Words were then stemmed, i.e., reducing inflected or sometimes derived words to their word stem, base, or root form (e.g., the words *fishing*, *fished*, and *fisher* to the stem *fish*). Additionally, due to sampling strategy, all words used in the initial API Twitter search (those listed in Supplementary Table 1), were also removed from the corpus, as their high frequency was inevitable.

Tokenizing at the word level can help greater understanding of the most frequently used words. However, examining different units of text (e.g., consecutive words) may provide a greater understanding of topics discussed. As such, we considered both single words (or tokens) and bigrams (two consecutive words) for both sector and/or phase. Additionally, bigrams provide further context in sentiment analysis, as it allows for the quantification of words preceded by negation. Negating words incorporated in analyses include: “aren’t”, “can’t”, “didn’t”, “don’t”, “hadn’t”, “hasn’t”, “no”, “not”, “shouldn’t”, “wasn’t”, “without”, and “won’t”. This highlights tokens which should be considered with caution. Pearson correlation coefficient was applied to test correlation of word use between phases and sectors.

Weighted log odds ratios

Log odds ratios were used to compare word usage between sectors and/or phases. The outcome, two “lexical histograms”, are taken from two sources X and Y, whose patterns of usage are contrasted. This method takes into account the likely sampling error in counts, discounting differences that are probably accidental, and enhances differences that are genuinely unexpected given the null hypothesis that both X and Y are making random selections from the same vocabulary. These features enabled differences in very frequent words to be detected.

Term frequency—Inverse document frequency

TF-IDF is a numerical statistic that evaluates how relevant a word is to a document in a collection of documents. This is done by multiplying two metrics: how many times a word appears in a document, and the inverse document frequency of the word across a set of documents. The logic of TF-IDF is that the words containing the greatest information about a particular document are the words that appear many times in that document, but in relatively few others. This was used to identify distinctive word use within a given sector and/or phase.

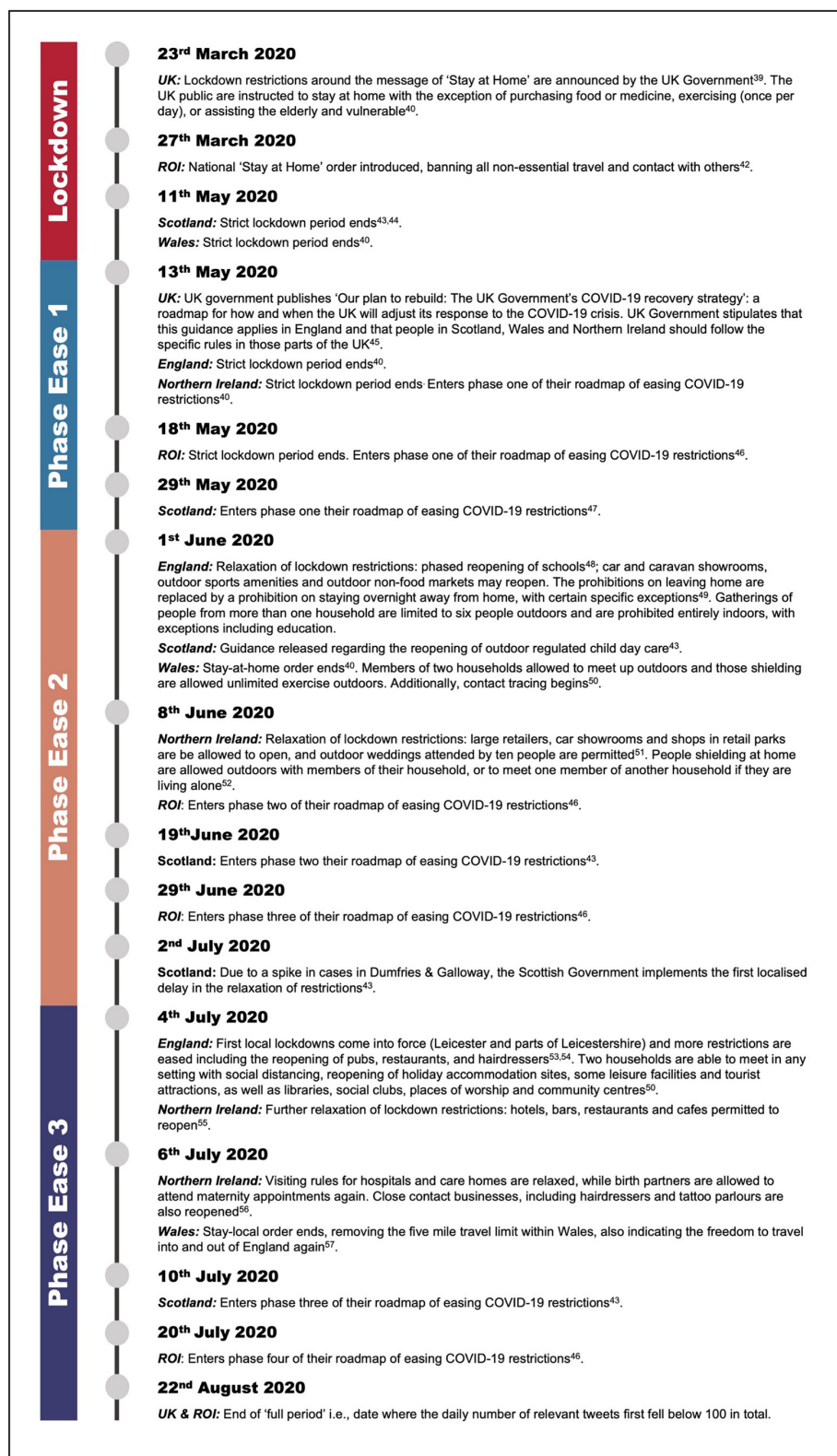
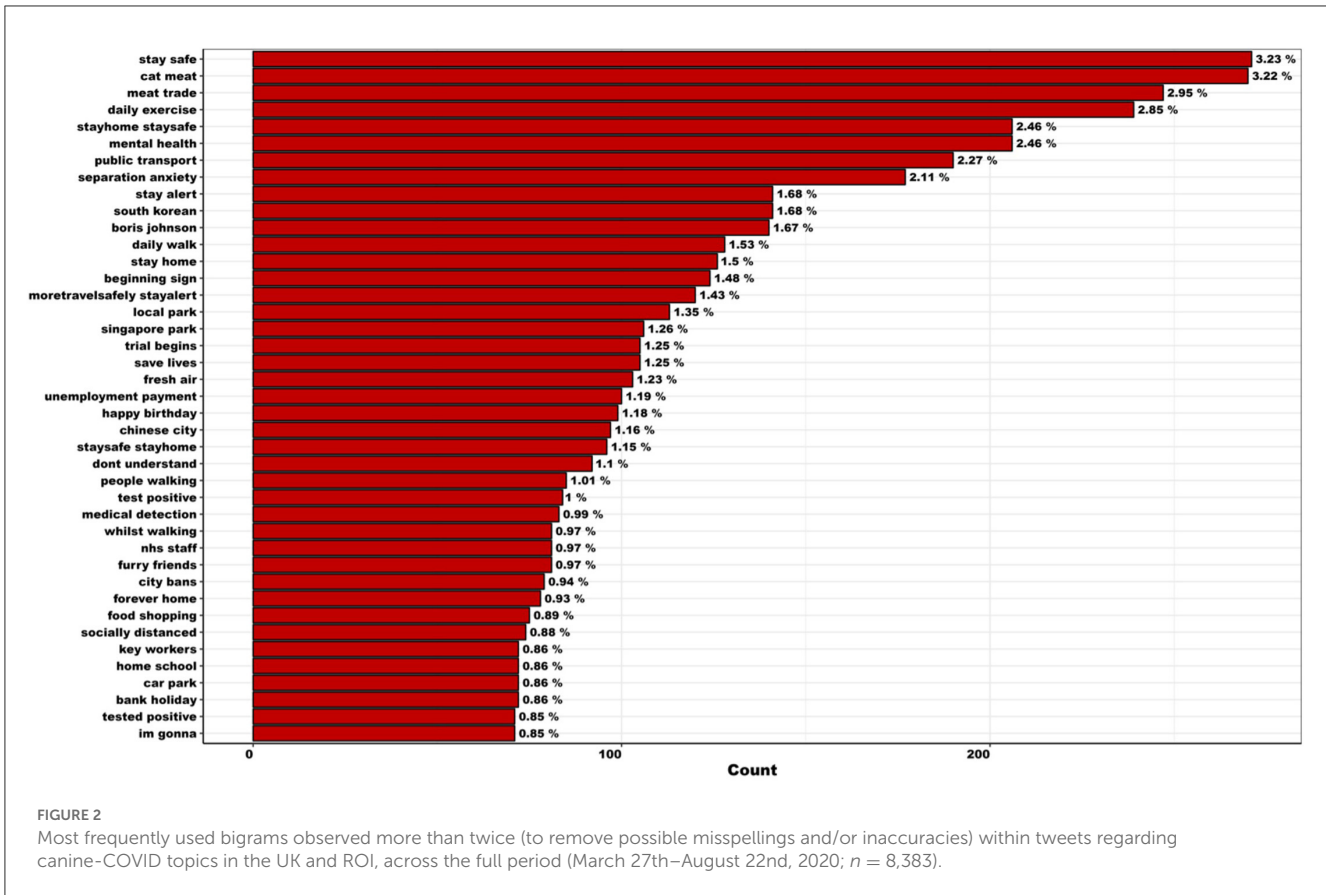


FIGURE 1

Dates of major events occurring across the UK nations (England, Northern Ireland, Wales, and Scotland) and Republic of Ireland (ROI), during the full period (March 27th–August 22nd, 2020), which are relevant to determining the four pandemic phases [Lockdown, Phase Ease 1, Phase Ease 2, and Phase Ease (39, 40, 42–57)].



Sentiment analysis

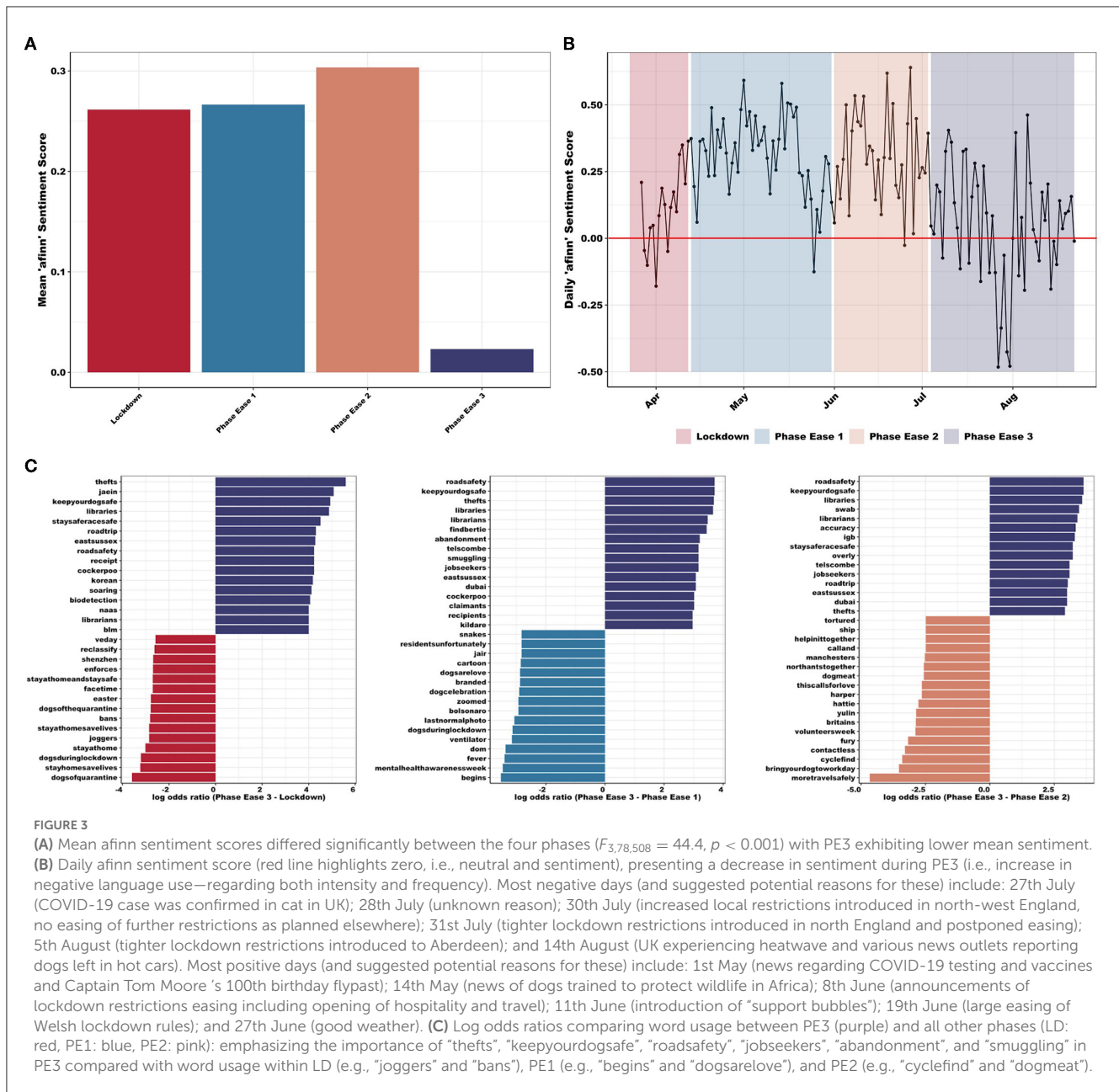
Sentiment analysis research has been accelerated with the development of several lexical resources. A wide variety of methods and dictionaries exist for evaluating the opinion or emotion in text. However, the aforementioned *tidytext* package provides access to several sentiment lexicons, including three general-purpose lexicons: *afinn* (60), *nrc* [syuzhet (61)], and *bing* (62). All three lexicons contain many English words and are based on tokens, i.e., single words. These words are assigned scores for positive/negative sentiment. The *afinn* lexicon assigns words with a score of -5 to 5 , with negative scores indicating negative sentiment and positive scores indicating positive sentiment. The *nrc* lexicon categorizes words into the following: positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The *bing* lexicon categorizes words in a binary fashion, into positive and negative categories. Due to the variation in scoring, all three lexicons were implemented in this project, to maximize opportunity to explore variation in measures of sentiment temporally and between sectors. Variations in sentiment were compared between sector and/or phase using univariate ANOVA. Any significant results from ANOVA testing were further tested using Tukey *post-hoc* test to determine which categorical groupings were different from the others.

Cross correlation functions and lagged regressions

In order to assess the relationship between two time series (y_t and x_t), the series y_t may be related to past lags of the x -series. To do so, the cross-correlation function [using function *CCF*, *tseries* package (63)] was implemented to determine the correlation between pair wise time series, as a function of the time lag or delay (64). This produces correlation coefficients between two time-series data at each lag. Consequently, we identified sectors which may influence others by assessing lags of mean sentiment within one sector that may be useful predictors of another. Relations between two sectors were then analyzed by means of Pearson correlations [using the function *cor.test*, *stats* package (30)], including the lagged period and correlations displayed using *DiagrammeR* package (65).

Ethics

Ethical approval for this study was granted by Dogs Trust Ethical Review Board (Reference Number: ERB036).



Results

Between March 27th and August 22nd, 2020, 73,551 tweets regarding COVID-19 and dogs were collected from the UK and ROI. During data cleaning and manipulation, 8,300 duplicate tweets were removed and 4,163 were identified as irrelevant. Consequently, the final dataset included 61,088 unique tweets, posted from 42,403 Twitter accounts. Of these 61,088 unique tweets, 75.1% ($n = 45,853$) were original and 24.9% ($n = 15,235$) were replies. Our analysis indicated a rapid increase and a slow decline in the volume of social media conversations regarding COVID-19 and dogs (Supplementary Figure 2). Mean number of relevant words per tweet was 1.9 (SE = 0.01, range = 1–23) and daily mean number of tweets was 410.0 (SE = 21.0, range = 101–1,036). The mean number of tweets posted per

account equated to 1.4 (SE = 0.01, range 1–391), and Twitter users applied 16,886 unique hashtags to original tweets. In total, 8,383 bigram combinations were observed more than twice (to remove possible misspellings and/or inaccuracies), throughout the full period, which are listed in Figure 2. The top 50 frequently used hashtags in original tweets are listed in Supplementary Figure 3.

Variation between phases

Summary statistics regarding tweets per phase are listed in Supplementary Table 3. Across the full period, "walk" remained the most frequently used token. While "distance", "home", "like", "love", "work", "help", and "need" remained consistently well used

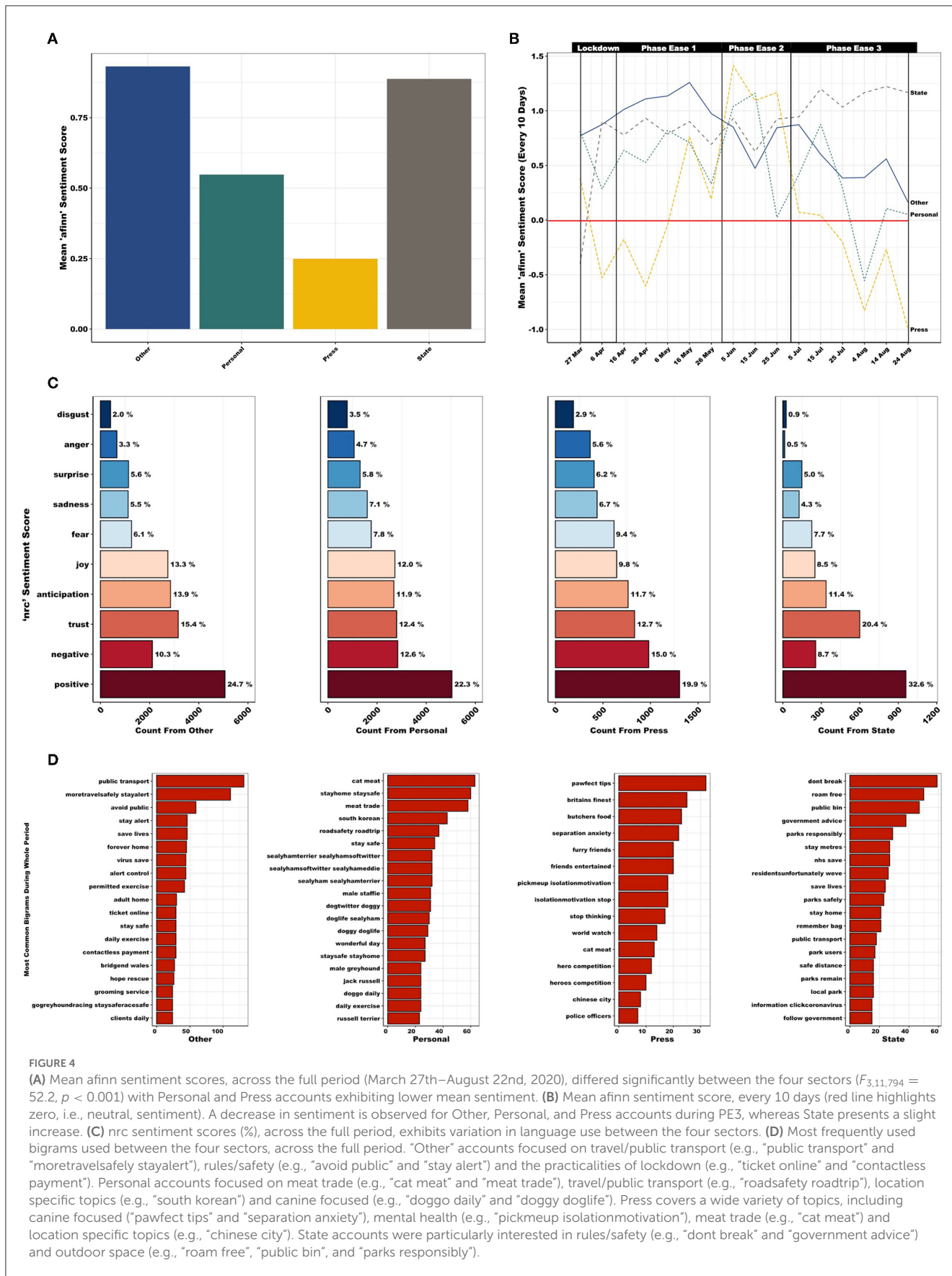
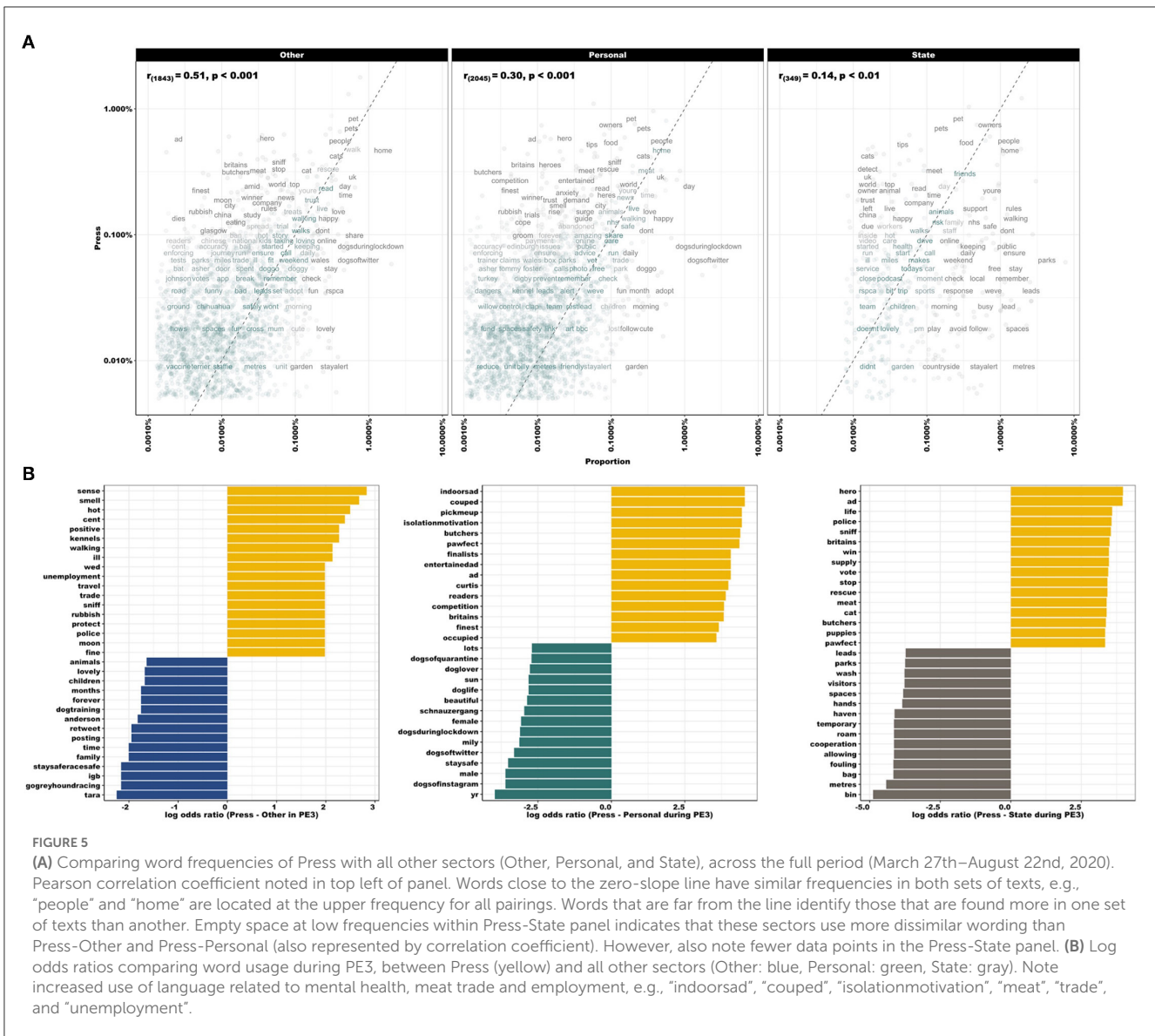


FIGURE 4

(A) Mean 'afinn' sentiment scores, across the full period (March 27th–August 22nd, 2020), differed significantly between the four sectors ($F_{3,11,794} = 52.2, p < 0.001$) with Personal and Press accounts exhibiting lower mean sentiment. (B) Mean 'afinn' sentiment score, every 10 days (red line highlights zero, i.e., neutral, sentiment). A decrease in sentiment is observed for Other, Personal, and Press accounts during PE3, whereas State presents a slight increase. (C) 'nrc' sentiment scores (%), across the full period, exhibits variation in language use between the four sectors. (D) Most frequently used bigrams used between the four sectors, across the full period. "Other" accounts focused on travel/public transport (e.g., "public transport" and "moretravelsafely stayalert"), rules/safety (e.g., "avoid public" and "stay alert") and the practicalities of lockdown (e.g., "ticket online" and "contactless payment"). Personal accounts focused on meat trade (e.g., "cat meat" and "meat trade"), travel/public transport (e.g., "roadsafety roadtrip"), location specific topics (e.g., "south korean") and canine focused (e.g., "doggo daily" and "doggy doglife"). Press covers a wide variety of topics, including canine focused ("pawfect tips" and "separation anxiety"), mental health (e.g., "pickmeup isolationmotivation"), meat trade (e.g., "cat meat") and location specific topics (e.g., "chinese city"). State accounts were particularly interested in rules/safety (e.g., "dont break" and "government advice") and outdoor space (e.g., "roam free", "public bin", and "parks responsibly").



(among others), the popularity of the initially well used hashtags “dogsduringlockdown” and “dogsoftwitter” diminished over time (Supplementary Figure 4). It is important to note that “help”, “like”, and “love” were frequently preceded by a negation term (e.g., “didn’t”, “don’t”, “no”, “not”, “wont”; Supplementary Figure 5). Consequently, these tokens should be considered with caution. Frequently used bigrams used throughout the full period, include meat trade and separation anxiety (Supplementary Figure 6).

Overall, the average sentiment score decreased significantly in PE3, with respect to *afinn* sentiment scores (Figures 3A, B). Mean *afinn* sentiment score differed significantly between the four phases ($F_{3,78,508} = 44.4, p < 0.001$; Figure 3A) with PE3 exhibiting lower mean sentiment score [mean (SE) = 0.02 (0.02)] than LD [mean (SE) = 0.26 (0.01); $p < 0.001$], PE1 [mean (SE) = 0.27 (0.02); $p < 0.001$], or PE2 [mean (SE) = 0.30 (0.02); $p < 0.001$]. Examination of daily sentiment scores across the full period identified the most negative days falling within PE3 (Figure 3B). PE3 also exhibited increased language related to “anger”, “sadness”,

“fear”, and “negative” sentiment, while also showing a decrease in “joy” and “positivity” (*nrc* sentiment; Supplementary Figure 7). Furthermore, PE3 presented lower total cumulative and mean *bing* sentiment score (Supplementary Figure 8).

As PE3 exhibited a significantly lower mean sentiment score (Figure 3A), word frequencies were compared between PE3 and all others (LD, PE1, and PE2). Due to the large number of tweets and high frequency of common words, correlation between all three remained consistently high (Supplementary Figure 9). Words with similar high frequencies in both sets of texts included “people”, “home”, “walk”, and “dogsoftwitter”. PE3-LD comparison presents “abandonment” and “unemployment” commonly within PE3 tweets, while “stayhomesaveslives” and “selfisolation” are frequently found in LD [$r_{(11,873)} = 0.86, p < 0.001$]. When comparing PE3-PE1, the latter commonly uses words such as “begins” and “lockdownlife”, whereas “staff” and “muzzle” are used more often in PE3 [$r_{(9,030)} = 0.88, p < 0.001$]. During PE2, words such as “transport” and “cycle”

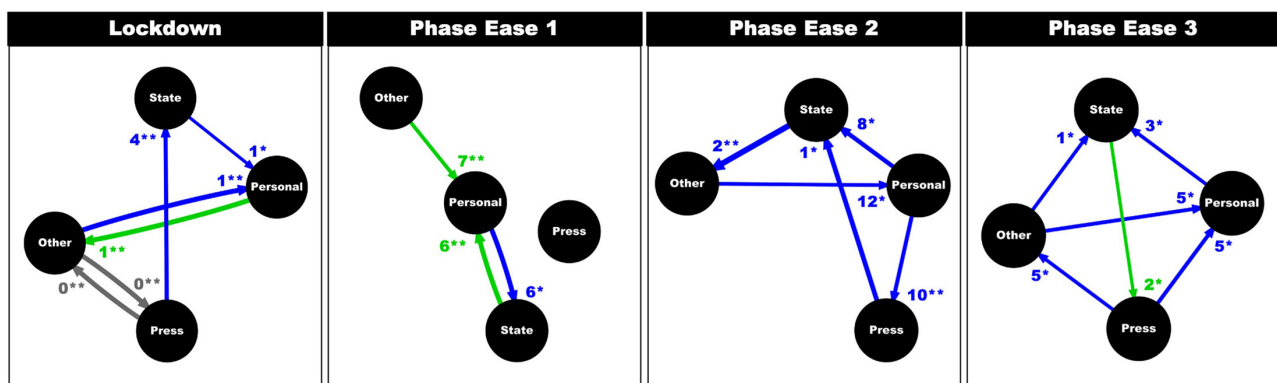


FIGURE 6

Results from cross correlation functions and lagged regressions, per phase. Each flowchart exhibits relationships between the four sectors, with regards to positive (green; above average sentiment in X is likely to lead to above average sentiment in Y), negative (blue; above average sentiment in X is likely to lead to below average sentiment in Y), or no (gray) correlation in mean sentiment. Arrows represent flow of influence and thickness of arrow represents strength of correlation. Lag (days) at which cross correlation is maximum is noted at arrowhead, along with statistical significance ($p < 0.5^*$, $p < 0.01^{**}$). Note that at the start of the pandemic (LD and PE1) Personal accounts were initially influenced by State and Other accounts (in both positive and negative directions). However, by PE3, Personal accounts were no longer correlated with State and Other accounts but Press and Other accounts.

are commonly located, whereas “thefts” and “hero” are more commonly found in PE3 [$r_{(8,929)} = 0.93$, $p < 0.001$]. Words also extend to lower frequencies for PE3-LD comparison, which indicates that PE3 and LD use more similar wording than PE3-PE1 or PE3-PE2. These results are mirrored in Figure 3C where words from LD, PE1, and PE2 are compared to PE3. These results emphasize the importance of “thefts”, “keepyourdogsafe”, “roadsafety”, “jobseekers”, “abandonment”, and “smuggling” in PE3 compared with all other phases, which are also reflected in the TF-IDF analyses (Supplementary Figure 10).

Variation between sectors

Across the full period, 386 “Personal” accounts, 104 “Press” accounts, 35 “State” accounts and 229 “Other” accounts contributed 9,259 tweets. Of these 754 sector-assigned accounts, 65 (i.e., 8.6%) were identified as potential bot accounts. However, these accounts were not removed from further analyses as verification required in-depth investigation. Of the 9,259 tweets assigned to the aforementioned sectors, 55.5% ($n = 5,135$) occurred during LD, 16.4% ($n = 1,515$) during PE1, 15.5% ($n = 1,438$) during PE2 and 12.6% ($n = 1,171$) during PE3 (Supplementary Figure 11). Mean number of relevant words per tweet was 2.6 (SE = 0.02, range = 1–19). Summary statistics regarding these tweets are listed in Supplementary Table 4.

Mean *afinn* sentiment score differed significantly between the four sectors ($F_{3,11,794} = 52.2$, $p < 0.001$; Figure 4A) with Press exhibiting lower mean sentiment score [mean (SE) = 0.25 (0.06)] than Other [mean (SE) = 0.93 (0.03); $p < 0.001$], Personal [mean (SE) = 0.55 (0.03); $p < 0.001$] and State [mean (SE) = 0.89 (0.05); $p < 0.001$]. Personal accounts also exhibited a lower mean sentiment score than Other ($p < 0.001$) and State ($p < 0.001$). Finer temporal examination of sentiment across the full period (i.e., mean sentiment score every 10 days; Figure 4B) shows that Other,

Personal and Press decrease in sentiment during PE3, whereas State shows a slight increase. With regards to *nrc* sentiment (Figure 4C), Other accounts score highly for “joy” and “anticipation”, while scoring low for “fear”. Personal accounts score highly for “disgust” and “sadness”, while scoring low for “trust”. Press accounts score poorly for “joy” and “positive”, and score highly for “anger”, “fear”, and “negative”. State accounts score highly for “trust” and “positive”, and score poorly for “disgust”, “anger”, and “negative”. Furthermore, Press accounts presented a lower total cumulative and mean *bing* sentiment score (Supplementary Figure 12).

Throughout the full period, “walk” and “help” remain consistently well used across all sectors (Supplementary Figure 13). However, popular hashtags, e.g., “dogsduringlockdown” and “dogsoftwitter”, and more emotive language, e.g., “love” and “like” are only present within Other and Personal accounts (however, refer to Supplementary Figure 5 regarding negation issues). Press accounts frequently use terminology related to “pet” and “owner”, e.g., “train”, “warn”, “food”, and “tip”. Meanwhile State accounts refer to restrictions and public spaces often, e.g., “park”, “distanc-”, “bin”, “rule” etc. The most frequent bigrams per sector suggests that the main foci of conversations varied between sectors (Figure 4D). “Other” accounts focused on travel/public transport, rules/safety, and the practicalities of lockdown. Personal accounts focused on meat trade, travel/public transport, location specific and canine focused topics. Press covered a wide variety of topics, including canine focused, mental health, meat trade and location specific topics. State accounts were particularly interested in rules/safety and outdoor space.

Figure 5A compares word frequencies of Press with all other disciplines (Other, Personal and State). In the Press-Other panel, “dies” and “finest” are found in Press tweets, while “lovely” and “share” are frequently found in Other [$r_{(1,843)} = 0.51$, $p < 0.001$]. When comparing Press-Personal, the latter used words such as “cute” and “adopt”, whereas Press commonly used “hero” and “anxiety” [$r_{(2,045)} = 0.30$, $p < 0.001$]. Within State

accounts, words such as “park” and “spaces” were commonly found, whereas “cats” and “owner” were more commonly found in Press tweets [$r_{(349)} = 0.14$, $p < 0.01$]. Press-Other sectors were most highly correlated, whilst Press-State use more dissimilar wording than Press-Other and Press-Personal, indicated by the empty space at low frequencies within the Press-State panel. Log odds ratios comparing word usage during PE3 between Press and all other sectors (Figure 5B) emphasized the importance of “indoorsad”, “unemployment”, “meat”, “trade”, “couped”, and “isolationmotivation” within Press accounts, compared with word usage in Other (e.g., “gogreyhoundracing”, “family”, and “dogtraining”), Personal (e.g., “dogsofinstagram”, “beautiful”, and “doglover”) or State (e.g., “bin”, “fouling”, and “parks”) accounts.

Cross correlation functions and lagged regressions

During LD, above average sentiment (i.e., positive language use) from Other and State accounts correlated with below average sentiment (i.e., greater use of negative language) from Personal accounts, on the following day [CCF(-1) = 0.40, $p > 0.01$; CCF(-1) = 0.26, $p = 0.05$; Figure 6]. Simultaneously, above average sentiment in Personal accounts was followed by increased sentiment in Other accounts 1 day later [CCF(1) = 0.40; $p < 0.01$]. Negative language use (i.e., below average sentiment) from Press accounts correlated with an increase in positive language use (i.e., above average sentiment) by State accounts, 4 days later [CCF(-4) = 0.34, $p < 0.01$]. Press and Other exhibited similar sentiment on the same day [CCF(0) = 0.39; $p < 0.01$ for both].

During PE1, sentiment with Other and State was still associated with that of Personal accounts. However, during this phase, unlike above, these relationships were positive: above average sentiment in Other and State was followed by above average sentiment within Personal accounts 7 and 6 days later, respectively [CCF(7) = 0.30; $p = 0.01$; CCF(6) = 0.39, $p < 0.01$]. Additionally, negative language use (i.e., below average sentiment) in Personal accounts was associated with an increase of positive language (i.e., above average sentiment) within State accounts 6 days later [CCF(-6) = 0.39; $p = 0.03$].

During PE2, above average sentiment within Other accounts were likely to precede below average sentiment within Personal accounts 12 days later [CCF(-12) = 0.29, $p = 0.02$]. Below average sentiment in Personal accounts preceded increased sentiment in Press and State accounts 10 and 8 days later [CCF(-10) = 0.32; $p < 0.01$; CCF(-8) = 0.35; $p = 0.02$]. Below average sentiment score in Press was associated with above average sentiment score in State 1 day later [CCF(-1) = 0.38, $p = 0.02$], and below average sentiment score in State was associated with above average sentiment score in Other 2 days later [CCF(-2) = 0.46, $p < 0.01$].

During PE3, above average sentiment in Other and Press were likely to be followed by below average sentiment within Personal accounts 5 days later [CCF(-5) = 0.30; $p = 0.03$; CCF(-5) = 0.31; $p = 0.02$]. Below average sentiment in Personal accounts preceded above average sentiment scores within State accounts 3 days later [CCF(-3) = 0.28; $p < 0.05$]. Below average sentiment score in Press was associated with above average sentiment score in

Other 5 days later [CCF(-5) = 0.28; $p = 0.04$]. Meanwhile, above average sentiment score in State was associated with above average sentiment score in Press 2 day later [CCF(2) = 0.26; $p = 0.02$], and above average sentiment score in Other was associated with below average sentiment score in State 1 day later [CCF(-1) = 0.28; $p = 0.04$].

Thus, during LD and PE1, Personal accounts were influenced by State and Other accounts, initially negatively (1 day later) and then positively (~1 week later), respectively. However, by PE3, whilst Personal accounts were still influenced by Other (5 days later), they were no longer influenced by State. Instead, sentiment displayed by Personal accounts was more closely aligned with Press accounts, 5 days later. Furthermore, Personal accounts were only positively influenced during PE1, by Other and State accounts (~1 week later). During all other phases (LD, PE2, and PE3), Personal accounts were negatively influenced by one or more sector. Finally, lag (days) at which cross correlation was maximum, seems to increase as time from the start date expands.

Discussion

We present temporal and sector specific insights into online behavior and cohort opinion, obtained through the application of text mining and sentiment analysis. We focus on canine welfare during a pandemic as an example topic, to demonstrate the functionality and effectiveness of applying these tools to scraped Twitter data. This paper aims to serve as proof of concept for applying these computational tools to topics related to animal welfare, in the hope of encouraging the application of these methodologies within the sector. However, please note limitations of the above methodologies are discussed in [Supplementary Note 1](#).

Gathering insight into public opinion, along with tracking and assessing main topics discussed online, aids our understanding of attitudes. Attitudes enable humans to determine, often very quickly, who to interact with, which products to purchase and/or which behaviors to engage in Sheeran et al. (66) and Maio and Olson (67). Given that human behavior shapes the welfare of our companion animals, due to the nature of the pet-owner relationship, it is important that we understand the public's attitude toward animals and their welfare. Due to this, a key interest of this paper was to highlight the capability of using Twitter data to identify topical foci of conversations regarding canine-COVID topics in the UK and ROI.

Within our dataset, tweets regularly mentioned positive tokens such as “love” and “together”. Several studies have identified dogs as being a source of purpose, routine, and entertainment for their owners during lockdown (25–27, 68). Positive implications of the pandemic have included owners reporting improved emotional bonds and/or more regular interaction with their pets (25, 69, 70). It is therefore unsurprising that people's focus turned toward this bond. Dogs are often depicted as providing their owners with comfort and feelings of relief during uncertain times, by offering companionship and interaction that would otherwise be lacking (23, 68, 69). This sense of companionship may have influenced the increase in web interest regarding adoption of cats and dogs during the early phase of the COVID-19 pandemic (23), which has been suggested to be sustainable for cats, but not dogs (71). This

topic may be present within our dataset, due to the frequently used token “adopt” within Personal accounts, and “smuggling” within PE3 conversations, potentially referring to the increased demand for puppies and associated intensification of large-scale breeding and/or rise of illegal puppy smuggling.

While the pandemic provided a unique opportunity to build bonds between owner and pet, there is growing concern regarding the impact it may have had on canine welfare. For example, the most prominent token used throughout all phases was “walk”, along with associated bigrams: “local park”, “fresh air”, “whilst walking”, and “daily walk”. Due to lockdown restrictions and further guidelines (72–74), it is likely that there has been a change to daily routines, which may have altered the frequency, duration, and location that owners walk their dog. Research surveying dog owners found that for many, dog walking became a treasured experience, offering the opportunity to get outdoors, maintain some form of routine and improve wellbeing (25). Despite this fact, others have reported that during lockdown, dogs were typically walked less often, for less time daily and had fewer opportunities to interact with other dogs: thus, decreasing opportunities for enrichment, socialization, and cognitive stimulation (24, 75). Either in response to, or predicting negative fallout as a result of this, Press accounts frequently referred to terminology relating to these issues, e.g., “pet”, “owner”, “train”, “warn”, “food”, and “tip”. Furthermore, due to local restrictions, reports of high numbers of dog walkers flocking to local outdoor spaces boomed (76), potentially increasing the amount of dog fouling in these spaces. Dog fouling presents a significant public health concern and can reduce the mental and physical wellbeing of nearby residents (77, 78). Furthermore, it has been stated that “dog excrement can have a significant economic impact in terms of deterring inward investment and tourism” to an area (79). Given these prior reasons, and the fact that activities associated with dog waste collection and disposal can be financially restrictive for local authorities, it is unsurprising that State accounts frequently use terms related to dog fouling, e.g., “public bin”, “remember bag”, and “parks responsibly”, along with “bin”, “bag”, and “fouling”. This topic is especially noteworthy as access to green space during the pandemic has been linked to a wide range of mental health outcomes and has been suggested as an essential quality-of-life element (80, 81).

While there were several foci of conversation, two topics remained consistently prominent throughout the full period: meat trade and separation anxiety, particularly within Personal and Press accounts. Discussion of the meat trade mostly referred to international campaigns seeking world-wide bans of both dog and cat meat-trade, along with associated closures of wet markets. With regards to separation anxiety, as a result of lockdown measures, dogs are reported to have spent more time in the company of their owner (24). With the relaxation of these measures, resultant changes to dog management and time spent alone may have increased likelihood of longer-term welfare issues such as dogs displaying separation-related behaviors, e.g., excessive barking, aggression, and destructive behavior (24, 82). As behavioral problems are reported to be one of the main reasons for the relinquishment to shelters (83), there is growing concern that the pandemic has affected adoption and

abandonment of dogs (23). This concern is mirrored within our data by the frequent use of “abandonment” during PE3, and “adopt” within Personal accounts. While we are yet to see the full impact of the pandemic on relinquishment rates, our findings, and those from previous studies highlight the importance of further research into crisis-driven changes in human–animal relationships.

During PE3, issues surrounding dog theft became prominent, e.g., “keepyourdogsafe” and “thefts”. Results of a BBC freedom of information request stated that five policing areas recorded a double-digit increase in the total number of dog thefts reported between January and July 2020, compared with the previous year (84). Furthermore, there were an estimated 2,000 incidents of dog theft reported in England and Wales in 2020 (85). This apparent increase in dog theft may be associated with the increasing demand for dogs during lockdown (23, 71). However, due to a lack of informative, comparable, and accessible datasets regarding dog thefts, analyzing spatiotemporal patterns, including incidence, remains very challenging (86). Future amendments to sentencing guidelines (85, 87) associated with the Theft Act (1968) (88) should consider adopting a standardized, centrally held data management system with a robust identifier for “pet theft”, allowing for a stronger spatial and temporal evidence base regarding the problem. The urgency of this topic, within the public domain, is clear within our results.

Hashtags can be viewed as topical markers, an indication of the tweet context or as the core topic expressed in the tweet. As such, researchers often assume that relevant populations consist of Twitter users who index their tweets with specific hashtags. Within this data, while the most frequently used hashtags “dogsduringlockdown” and “dogsoftwitter” appeared frequently within Personal and Other accounts, Press and State did not commonly utilize trending hashtags. Given the potential to improve one’s “searchability” and attract a relevant audience by utilizing clear hashtags (89, 90), it may be prudent to suggest this to be a missed opportunity. However, while these hashtags were utilized frequently during LD and PE1, their popularity diminished with time, reflecting the fast-paced and evolving nature of Twitter. Romero et al. (91) stated that different categories of hashtags have different propagation patterns, introducing the distinction between “stickiness” and “persistence”: arguing that some classes of hashtag are more persistent than others. Our work shares these observations and highlights that delineating populations *via* hashtags may create analytical issues for researchers downstream, i.e., when tweets must be categorized as relevant vs. “noise”, as hashtag use are both sector and temporally biased.

Studies have reported that people living under lockdown measures are more prone to evolve various psychological symptoms, e.g., stress, depression, emotional fatigue, insomnia, and signs of post-traumatic anxiety (92). This may be reflected in the frequent use of terms relating to mental health throughout the full period, but especially during PE3, i.e., “indoorsad”, “couped”, and “isolationmotivation”. Thus, it is somewhat unsurprising that the psychological stress of the pandemic generated a stark negative response during PE3 within Personal, Press and Other

accounts, represented by the rapid rise in negative sentiment, i.e., language related to “anger”, “sadness”, “fear”, and “negativity”. These results are supported by a previous study focusing on Chinese social media usage, which found an increase in negative emotions (i.e., depression and anxiety) and a decrease in positive emotions and life satisfaction when compared to pre COVID-19 times (93). In contrast, State accounts presented a slight increase in positive sentiment during PE3, frequently using language relating to “trust” and “positivity” and rarely using language relating to “disgust”, “anger”, “sadness”, or “negativity”. There are multiple factors that may influence variation in the language used by sectors online, e.g., target audience, communication direction flow, marketing strategies (reactive vs. proactive) and personal/corporation aims. However, the temporal differences in sentiment portrayal evident across groups suggests influence between groups and, potentially, emotional contagion, i.e., whereby user sentiment may be affected by others (94). The impact of emotional contagion may be considered positive (e.g., informed science communication) or negative (e.g., inaccurate information exchange leading to detrimental attitudes and diminished animal welfare). We examined the influence of one sector upon another and suggested the “electronic word of mouth” time delay that may exist between two sectors. Early in the pandemic (LD and PE1) Personal accounts appeared to be initially influenced by State and Other accounts. However, by PE3, Personal accounts were influenced by Press and Other accounts. Consequently, there was evidence of a shift in sector influence on public perception (at least regarding canine-COVID topics). These results suggest that the most influential sources, toward the public, may change over time: in this case State becoming less influential and Press more. This shift in public responsiveness to government messaging during times of crises has been previously suggested (95, 96). Additionally, the number of days taken to see change in sentiment within the affected sector may vary over time (in this case study it appeared to increase). Thus, influence between sectors may be more rapidly evident closer to the original starting point.

Effective and temporally accurate data is essential to the successful management and completion of any corporation or charity aims. Whilst these methods initially require computational infrastructure, knowledge and training, the sourcing of data is rapid, robust, rich, and inexpensive, especially in comparison to alternative methods aiming to ascertain attitudes, such as focus groups and online questionnaires. Given the real-time streaming nature of Twitter, ability to search archives for relevant topics, open access statistical products, free data visualization platforms, and online help communities, the methods described here provide a unique, yet untapped, tool for the animal welfare sector. Furthermore, while keeping pace with the online community may prove challenging, doing so may be beneficial to multiple stakeholders, as it should increase their online presence, embed them within pertinent conversations, increase their likelihood of remaining relevant and provide insights into online behavior and cohort opinion.

Our findings do not elucidate the role that sectors have on public perception of specific canine-COVID topics, we simply

highlight that the information conveyed, and language used by sectors can have an impact on public perception. Sector influence may guide public attitudes toward animal welfare, which may, in turn, influence behavior (94). While this could directly impact upon human-animal interactions, both negatively and positively: it could also indirectly affect animal welfare by altering support toward specific topics, e.g., political causes. This highlights the importance of sector responsibility regarding appropriate and accurate messaging. As sectors were found to display varying sentiment, it is paramount that we continue to examine the influence of certain groups on public opinion. We suggest future studies elucidate opinion and sentiment surrounding specific animal welfare topics, potentially before and after behavioral intervention campaigns, to better inform the development of proactive response strategies.

Data availability statement

The datasets presented in this article are not readily available due to Twitter’s Developer Agreement. Requests to access the datasets should be directed to kirsten.mcmillan@dogstrut.org.uk.

Ethics statement

Ethical approval for this study was granted by Dogs Trust Ethical Review Board (Reference Number: ERB036). Written informed consent from the participants’ legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

RC collected the data. KM designed methodology, analyzed the data, and led the writing of the manuscript. RC, KM, and KA conceived the ideas, contributed critically to the drafts, and gave final approval for publication. All authors contributed to the article and approved the submitted version.

Funding

All data collection and analyses were supported by Dogs Trust.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2023.1074542/full#supplementary-material>

References

- Statista. *Number of Monthly Active Twitter Users Worldwide from 1st Quarter 2010 to 1st Quarter 2019*. (2021). Available online at: <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/> (accessed January 20, 2021).
- Muñoz-Expósito M, Oviedo-García MÁ, Castellanos-Verdugo M. How to measure engagement in Twitter: advancing a metric. *Internet Res.* (2017) 27:1122–48. doi: 10.1108/IntR-06-2016-0170
- Irfan R, King CK, Grages D, Ewen S, Khan SU, Madani SA, et al. A survey on text mining in social networks. *Knowl Eng Rev.* (2015) 30:157–70. doi: 10.1017/S0269888914000277
- Matheson NJ, Lehner PJ. How does SARS-CoV-2 cause COVID-19? *Science.* (2020) 369:510–1. doi: 10.1126/science.abc6156
- Miazga J, Hachaj T. Evaluation of most popular sentiment lexicons coverage on various datasets. In: *Proceedings of the 2019 2nd International Conference on Sensors, Signal and Image Processing* Prague. (2019). p. 86–90.
- Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. *Sci Rep.* (2020) 10:1–10. doi: 10.1038/s41598-020-73510-5
- van Vliet L, Törnberg P, Uitermark J. The twitter parliamentarian database: analyzing twitter politics across 26 countries. *PLoS ONE.* (2020) 15:e0237073. doi: 10.1371/journal.pone.0237073
- Matalon Y, Magdaci O, Almozilino A, Yamin D. Using sentiment analysis to predict opinion inversion in Tweets of political communication. *Sci Rep.* (2021) 11:1–9. doi: 10.1038/s41598-021-86510-w
- Crooks A, Croitoru A, Stefanidis A, Radzikowski J. #Earthquake: Twitter as a distributed sensor system. *Trans GIS.* (2013) 17:124–47. doi: 10.1111/j.1467-9671.2012.01359.x
- Zou L, Lam NS, Cai H, Qiang Y. Mining Twitter data for improved understanding of disaster resilience. *Ann Am Assoc Geogr.* (2018) 108:1422–41. doi: 10.1080/24694452.2017.1421897
- Daume S. Mining Twitter to monitor invasive alien species - an analytical framework and sample information topologies. *Ecol Inform.* (2016) 31:70–82. doi: 10.1016/j.ecoinf.2015.11.014
- Hart AG, Carpenter WS, Hlustik-Smith E, Reed M, Goodenough AE. Testing the potential of Twitter mining methods for data acquisition: evaluating novel opportunities for ecological research in multiple taxa. *Methods Ecol Evol.* (2018) 9:2194–205. doi: 10.1111/2041-210X.13063
- Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EHY, Olsen JM et al. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. *PLoS ONE.* (2015) 10:e0139701. doi: 10.1371/journal.pone.0139701
- Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: Sentiment analysis and topic modeling study. *JMIR Public Health Surveill.* (2020) 6:e21978. doi: 10.2196/21978
- Buddle EA, Bray HJ, Pitchford WS. Keeping it 'inside the fence': an examination of responses to a farm-animal welfare issue on Twitter. *Anim Prod Sci.* (2018) 58:435–44. doi: 10.1071/AN16634
- Wonneberger A, Hellsten IR, Jacobs SHJ. Hashtag activism and the configuration of counterpublics: Dutch animal welfare debates on Twitter. *Inform Commun Soc.* (2021) 24:1694–711. doi: 10.1080/1369118X.2020.1720770
- Propper C, Stoye G, Zaranko B. The wider impacts of the coronavirus pandemic on the NHS. *Fiscal Studies.* (2020) 41:227. doi: 10.1111/1475-5890.12227
- Corlett RT, Primack RB, Devictor V, Maas B, Goswami VR, Bates AE, et al. Impacts of the coronavirus pandemic on biodiversity conservation. *Biol Conserv.* (2020) 246:108571. doi: 10.1016/j.biocon.2020.108571
- Gortázar C, de la Fuente J. COVID-19 is likely to impact animal health. *Prev Vet Med.* (2020) 180:105030. doi: 10.1016/j.prevetmed.2020.105030
- Neupane D. How conservation will be impacted in the COVID-19 pandemic. *Wildlife Biol.* (2020) 2:727. doi: 10.2981/wlb.00727
- Van der Waal K, Deen J. Global trends in infectious diseases of swine. *Proc Natl Acad Sci.* (2018) 115:11495–500. doi: 10.1073/pnas.1806068115
- Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science.* (2020) 368:1016–20. doi: 10.1126/science.abb7015
- Morgan L, Protopopova A, Birkler RI, Itin-Shwartz B, Sutton GA, Gamliel A, et al. Human–dog relationships during the COVID-19 pandemic: booming dog adoption during social isolation. *Humanit Soc Sci Commun.* (2020) 7:1–11. doi: 10.1057/s41599-020-00649-x
- Christley RM, Murray JK, Anderson KL, Buckland EL, Casey RA, Harvey ND, et al. Impact of the first COVID-19 lockdown on management of pet dogs in the UK. *Animals.* (2021) 11:5. doi: 10.3390/ani11010005
- Holland KE, Owczarczak-Garstecka SC, Anderson KL, Casey RA, Christley RM, Harris L, et al. "More attention than usual": a thematic analysis of dog ownership experiences in the UK during the first COVID-19 lockdown. *Animals.* (2021) 11:240. doi: 10.3390/ani11010240
- Shoesmith E, Santos de Assis L, Shahab L, Ratschen E, Toner P, Kale D, et al. The perceived impact of the first UK COVID-19 lockdown on companion animal welfare and behaviour: a mixed-method study of associations with owner mental health. *Int J Environ Res Public Health.* (2021) 18:6171. doi: 10.3390/ijerph18116171
- Shoesmith E, Santos de Assis L, Shahab L, Ratschen E, Toner P, Kale D, et al. The influence of human-animal interactions on mental and physical health during the first COVID-19 lockdown phase in the UK: a qualitative exploration. *Int J Environ Res Public Health.* (2021) 18:976. doi: 10.3390/ijerph18030976
- Kumar S, Morstatter F, Liu H. *Twitter Data Analytics*. New York, NY: Springer (2014).
- Kearney MW. *tweet: collecting and analyzing Twitter data.* *J Open Source Softw.* (2019) 4:1829. doi: 10.21105/joss.01829
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna (2020). Available online at: <https://www.R-project.org> (accessed March 15, 2020).
- Wickham H, Francois R, Henry L, Müller K. *Dplyr: A Grammar of Data Manipulation*. R package version 0.8.5 (2015). Available online at: <https://CRAN.R-project.org/package=dplyr> (accessed September 22, 2017).
- Wickham H. *Stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.4.0 (2015). Available online at: <https://CRAN.R-project.org/package=stringr> (accessed January 3, 2020).
- Wickham H. *Tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1 (2017). Available online at: <https://CRAN.R-project.org/package=tidyverse> (accessed June 20, 2018).
- Silge J, Robinson D. *tidytext: Text mining and analysis using tidy data principles in R.* *J Open Source Soft.* (2016) 1:37. doi: 10.21105/joss.00037
- Feinerer I. *Introduction to the tm Package Text Mining in R. R package version 3.1.* (2013). Available online at: <http://cran.r-project.org/web/packages/tm/vignettes/tm.pdf> (accessed June 20, 2018).
- Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag (2016).
- Fellows I. *wordcloud: Word Clouds. R Package Version 2.6.* (2018). Available online at: <https://CRAN.R-project.org/package=wordcloud> (accessed July 8, 2019).
- Luis von Ahn Research Group. *Offensive/Profane Word List.* (2018). Available online at: <https://www.cs.cmu.edu/~luis/bigblou/resources/bad-words.txt> (accessed June 1, 2019).
- GOV.UK. *Prime Minister's Statement on Coronavirus (COVID-19).* (2020). Available online at: <https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020> (accessed July 2, 2020).
- Cameron-Blake E, Tatlow H, Wood A, Hale T, Kira B, Petherick A, et al. *Variation in the Response to COVID-19 Across the Four Nations of the United Kingdom*. Oxford: Blavatnik School of Government, University of Oxford (2020).
- Sloan L, Morgan J. Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on twitter. *PLoS ONE.* (2015) 10:e0142209. doi: 10.1371/journal.pone.0142209

42. GOV.IE. *Public Health Measures in Place Until 12 April to Prevent Spreading COVID-19*. (2020). Available online at: <https://www.gov.ie/en/publication/3361b-public-health-updates/> (accessed January 2, 2022).
43. Scottish Parliament. *SPiCe Spotlight: Timeline of Coronavirus (COVID-19) in Scotland*. (2022). Available online at: <https://spice-spotlight.scot/2022/12/16/timeline-of-coronavirus-covid-19-in-scotland/> (accessed December 20, 2022).
44. Scottish Government. *Coronavirus (COVID-19): First Minister Address to the Nation*. (2020). Available online at: <https://www.gov.scot/publications/first-minister-address-nation-11-2020/> (accessed May 11, 2020).
45. GOV.UK. *Our Plan to Rebuild: The UK Government's Covid-19 Recovery Strategy*. (2020). Available online at: <https://www.gov.uk/government/publications/our-plan-to-rebuild-the-uk-governments-covid-19-recovery-strategy/our-plan-to-rebuild-the-uk-governments-covid-19-recovery-strategy> (accessed August 1, 2020).
46. GOV.IE. *Government Publishes Roadmap to Ease COVID-19 Restrictions and Reopen Ireland's Society and Economy*. (2020). Available online at: <https://www.gov.ie/en/press-release/e5e599-government-publishes-roadmap-to-ease-covid-19-restrictions-and-reopen/#phase-1-18-may> (accessed May 20, 2020).
47. Scottish Government. *Coronavirus (COVID-19) Update: First Minister's Speech*. (2020). Available online at: <https://www.gov.scot/publications/coronavirus-covid-19-update-first-ministers-speech-28-2020/> (accessed June 20, 2020).
48. Institute for Government. *Timeline of UK Government Coronavirus Lockdowns and Measures, March 2020 to December 2021*. (2021). Available online at: <https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns> (accessed January 3, 2022).
49. GOV.UK. *The Health Protection (Coronavirus, Restrictions) (England) (Amendment No. 3) Regulations (SI 558)*. (2020). Available online at: <https://www.legislation.gov.uk/ukksi/2020/558/made> (accessed January 5, 2022).
50. Welsh Parliament. *Coronavirus Timeline: Welsh and UK Government's Response: Research Briefing*. (2020). Available online at: <https://senedd.wales/media/fu0kdyg0/coronavirus-timeline-211220.pdf> (accessed February 20, 2020).
51. BBC News. *NI Aims to Allow Small Outdoor Weddings in June - 28 May*. (2020). Available online at: <https://www.bbc.co.uk/news/uk-northern-ireland-52809997> (accessed June 1, 2020).
52. BBC News. *People Shielding Allowed Outdoors from 8 June - 1 June*. (2020). Available online at: <https://www.bbc.co.uk/news/uk-northern-ireland-52874529> (accessed July 1, 2020).
53. Imperial College Covid-19 Response Team. *Impact of Non-pharmaceutical Interventions (NPIs) to Reduce Covid-19 Mortality and Healthcare Demand*. (2020). Available online at: <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf> (accessed June 30, 2020).
54. O'Hare R, van Elsland SL. *Coronavirus Measures May Have Already Averted up to 120,000 Deaths Across Europe*. Imperial College London (2020). Available online at: <https://www.imperial.ac.uk/news/196556/coronavirus-measures-have-already-averted-120000/> (accessed May 1, 2020).
55. BBC News. *NI's Hotels and Bars Reopen Their Doors - 3 July*. (2020). Available online at: <https://www.bbc.co.uk/news/uk-northern-ireland-53264251> (accessed July 20, 2020).
56. BBC News. *Hospital and Care Home Visits Resume in NI - 6 July*. (2020). Available online at: <https://www.bbc.co.uk/news/uk-northern-ireland-53293931> (accessed July 20, 2020).
57. Welsh Government. *Stay Local to be Lifted in Wales*. (2020). Available online at: <https://www.gov.wales/stay-local-to-be-lifted-in-wales> (accessed July 20, 2020).
58. Kearney MW. *Twitterbotnot: Detecting Twitter bots*. (2018). Available online at: <https://github.com/mkearney/twitterbotnot> (accessed June 1, 2019).
59. Chu Z, Gianvecchio S, Wang H, Jajodia S. Who is tweeting on Twitter: human, bot, or cyborg? In: *Proceedings of the 26th Annual Computer Security Applications Conference*. Austin, TX (2010). p. 21–30.
60. Bradley MM, Lang PJ. *Affective Norms for English Words (ANEW) Instruction Manual and Affective Ratings*. Technical Report C-1, the Center for Research in Psychophysiology University of Florida. (2009).
61. Mohammad SM, Turney PD. Crowdsourcing a wordsource association lexicon. *Comput Intell*. (2012) 29:436–65. doi: 10.1111/j.1467-8640.2012.00460.x
62. Liu B, Zhang L. A survey of opinion mining and sentiment analysis. In: C. Aggarwal, C. Zhai, editors. *Mining Text Data*. Boston, MA: Springer (2012). p. 415–63.
63. Trapletti A, Hornik K, LeBaron B, Hornik MK. *Package 'tseries'*. *R Package Version 0.10-47*. (2020). Available online at: <https://cran.r-project.org/web/packages/tseries/tseries.pdf> (accessed January 10, 2020).
64. Zhang L, Wu X. On the application of cross correlation function to subsample discrete time delay estimation. *Digit Signal Process*. (2006) 16:682–94. doi: 10.1016/j.dsp.2006.08.009
65. Iannone R. *Package 'DiagrammeR'*. *R Package Version 1.2.1*. (2016). Available online at: <https://cran.r-project.org/web/packages/DiagrammeR/DiagrammeR.pdf> (accessed January 9, 2020).
66. Sheeran P, Norman P, Orbell S. Evidence that intentions based on attitudes better predict behaviour than intentions based on subjective norms. *Eur J Soc Psychol*. (1999) 29:403–6. doi: 10.1002/(SICI)1099-0992(199903/05)29:2<403::AID-EJSP94>3.0.CO;2-A
67. Maio GR, Olson JM. Emergent themes and potential approaches to attitude function: the function-structure model of attitudes. In: Maio GR, Olson JM, editors. *Why We Evaluate: Functions of Attitudes*. New York, NY: Psychology Press (2000). p. 417–42.
68. Ratschen E, Shoesmith E, Shahab L, Silva K, Kale D, Toner P, et al. Human-animal relationships and interactions during the COVID-19 lockdown phase in the UK: investigating links with mental health and loneliness. *PLoS ONE*. (2020) 15:e0239397. doi: 10.1371/journal.pone.0239397
69. Bowen J, García E, Darder P, Argüelles J, Fatjó J. The effects of the Spanish COVID-19 lockdown on people, their pets and the human-animal bond. *J Vet Behav*. (2020) 40:75–91. doi: 10.1016/j.jveb.2020.05.013
70. Bussolari C, Currin-McCulloch J, Packman W, Kogan L, Erdman P. "I couldn't have asked for a better quarantine partner!": experiences with companion dogs during Covid-19. *Animals*. (2021) 11:330. doi: 10.3390/ani11020330
71. Ho J, Hussain S, Sparagano O. Did the COVID-19 pandemic spark a public interest in pet adoption? *Front Vet Sci*. (2021) 8:444. doi: 10.3389/fvets.2021.647308
72. UKGOV. *Staying at Home and Away from Others (Social Distancing)*. UKGOV: London, UK (2020). Available online at: <https://www.gov.uk/government/publications/full-guidance-on-staying-at-home-and-away-from-others> (accessed July 22, 2020).
73. Government of Ireland. *Daily Briefing on the Government's Response to COVID-19 - Friday 27 March 2020*. (2020). Available online at: <https://www.gov.ie/en/publication/aabc99-daily-briefing-on-the-governments-response-to-covid-19-friday-27-mar/> (accessed July 22, 2020).
74. Government of Ireland. *Public Health Measures in Place Until 5 May to Prevent Spreading COVID-19*. (2020). Available online at: <https://www.gov.ie/en/publication/cf9b0d-new-public-health-measures-effective-now-to-prevent-further-spread-o/> (accessed July 22, 2020).
75. Owczarczak-Garstecka SC, Graham TM, Archer DC, Westgarth C. Dog walking before and during the COVID-19 pandemic lockdown: experiences of UK dog owners. *Int J Environ Res and Public Health*. (2021) 18:6315. doi: 10.3390/ijerph18126315
76. Day BH. The value of greenspace under pandemic lockdown. *Environ Resour Econ*. (202) 76:1161–85. doi: 10.1007/s10640-020-00489-y
77. Mateus TL, Castro A, Ribeiro JN, Vieira-Pinto M. Multiple zoonotic parasites identified in dog feces collected in Ponte de Lima, Portugal - A potential threat to human health. *Int J Environ Res Public Health*. (2014) 11:9050–67. doi: 10.3390/ijerph110909050
78. Derges J, Lynch R, Chow A, Petticrew M, Draper A. Complaints about dog faeces as a symbolic representation of incivility in London, UK: a qualitative study. *Crit Public Health*. (2012) 22:419–25. doi: 10.1080/09581596.2012.710738
79. Atenstaedt RL, Jones S. Interventions to prevent dog fouling: a systematic review of evidence. *Public Health*. (2011) 125:90–2. doi: 10.1016/j.puhe.2010.09.006
80. Soga M, Evans MJ, Tsuchiya K, Fukano Y. A room with a green view: the importance of nearby nature for mental health during the COVID-19 pandemic. *Ecol Appl*. (2020) 31:e2248. doi: 10.1002/eap.2248
81. Kleinschroth F, Kowarik I. COVID-19 crisis demonstrates the urgent need for urban greenspaces. *Front Ecol Environ*. (2020) 18:318. doi: 10.1002/fee.2230
82. Flannigan G, Dodman NH. Risk factors and behaviors associated with separation anxiety in dogs. *J Am Vet Med Assoc*. (2001) 219:460–6. doi: 10.2460/javma.2001.219.460
83. Salman MD, Hutchison J, Ruch-Gallie R, Kogan L, New JC, Jr., Kass PH, et al. Behavioral reasons for relinquishment of dogs and cats to 12 shelters. *J Appl Animal Welfare Sci*. (2000) 3:93–106. doi: 10.1207/S15327604JAWS0302_2
84. Woodfield A. *Coronavirus: Lockdown Year 'Worst Ever' for Dog Thefts*. BBC News, East Midlands (2020). Available online at: <https://www.bbc.co.uk/news/uk-england-54372778> (accessed October 25, 2020).
85. Pet Theft Taskforce. *Policy Paper: Pet Theft Taskforce Report*. (2021). Available online at: <https://www.gov.uk/government/publications/pet-theft-taskforce-report/pet-theft-taskforce-report> (accessed September 26, 2021).
86. Allen D, Peacock A, Arathoon J. Spatialities of dog theft: a critical perspective. *Animals*. (2019) 9:209. doi: 10.3390/ani9050209
87. Sentencing Council. *Theft Offences: Definitive Guideline*. Sentencing Council (2016). Available online at: <https://www.sentencingcouncil.org.uk/publications/item/theft-offences-definitive-guideline/> (accessed October 20, 2019).
88. Wood JC. The theft act 1968. *Police J*. (1968) 41:551–7. doi: 10.1177/0032258X6804101203
89. Zappavigna M. Searchable talk: the linguistic functions of hashtags. *Soc Semiot*. (2015) 25:274–91. doi: 10.1080/10350330.2014.996948

90. Saxton GD, Niyirora J, Guo C, Waters R. #AdvocatingForChange: the strategic use of hashtags in social media advocacy. *Adv Soc Work.* (2015) 16:154–69. doi: 10.18060/17952
91. Romero DM, Meeder B, Kleinberg J. Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on twitter. In: *Proceedings of the 13th International Conference on World Wide Web.* New York, NY (2011).
92. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet.* (2020) 395:912–20. doi: 10.1016/S0140-6736(20)30460-8
93. Shigemura J, Ursano RJ, Morganstein JC, Kurosawa M, Benedek DM. Public responses to the novel 2019 coronavirus (2019-nCoV) in Japan: mental health consequences and target populations. *Psychiatry Clin Neurosci.* (2020) 74:281–2. doi: 10.1111/pcn.12988
94. Ferrara E, Yang Z. Measuring emotional contagion in social media. *PLoS ONE.* (2015) 10:e0142390. doi: 10.1371/journal.pone.0142390
95. Van der Weerd W, Timmermans DR, Beaujean DJ, Oudhoff J, Van Steenberghe JE. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in the Netherlands. *BMC Public Health.* (2011) 11:575. doi: 10.1186/1471-2458-11-575
96. Vardavas C, Odani S, Nikitara K, El Bahawi H, Kyriakos C, Taylor L, et al. Public perspective on the governmental response, communication and trust in the governmental decisions in mitigating COVID-19 early in the pandemic across the G7 countries. *Prev Med Rep.* (2021) 21:101252. doi: 10.1016/j.pmedr.2020.101252