Transfusion Medicine Reviews 36 (2022) 58–65

Contents lists available at ScienceDirect

Transfusion Medicine Reviews

journal homepage: https://www.journals.elsevier.com/transfusion-medicine-reviews

Blood Donation Narratives on Social Media: A Topic Modeling Study

Steven Ramondt, PhD[a,b,*], Peter Kerkhof[b], Eva-Maria Merz[a,c]

[a] Donor Studies, Department. of Donor Medicine Research, Sanquin Research, Amsterdam, the Netherlands
[b] Department of Communication Science, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[c] Center for Philanthropic Studies, Department of Sociology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Available online 17 October 2021

Keywords:
Donor research
Social media
Blood donor
Recruitment and retention
Facebook
Twitter
Structural topic modeling

ABSTRACT

Social media have shown great potential for producing significant changes in behavior and have become the cornerstone for many public health and agency efforts. The nonprofit sector -including blood collection agencies- has adopted social media to aid their cause and reach their goals. However, despite the tremendous impact of social media on society and its promising role for donor recruitment and retention, it has been overlooked in donor research. This study therefore sought to map the social media landscape around blood donation. We showcase an inductive computational method to make sense of vast amounts of dynamic unstructured blood donation text data that exists on social media. With this method, we display what is discussed about blood donation on social media, how these topics are distributed on Facebook and Twitter, and how the prevalence of these topics changes over time. We applied structural topic modeling on 7 years of Dutch blood donation Facebook and Twitter data by the general public. We found 25 topics clustered in 6 distinct clusters. Over time, there is a substantial reduction of messages in which donors announce their donations. Topics that emphasize the positives of blood donation, including donor identity-related topics, are rising. In addition, the findings show a clear social media platform contrast. Topics related to campaigns and controversial policies were found more on Twitter and positive donation topics and topics related to the donation process more on Facebook. To make optimal use of social media for recruitment and retention campaigns and efforts, blood collection agencies should recognize the turbulent environment in which they take place. Monitoring public opinions about blood donation will help blood collection agencies make strategic choices and utilize social media more effectively.

Introduction

As social media have become integrated into society, almost all nonprofits have embraced social media to reach large audiences [1]. Utilization of social media is high, with the majority (72%) of all American adults using social media, and 90% utilization in the 18-29 group [2]. Social media have shown great potential for producing significant changes in behavior and have become the cornerstone for many public health and agency efforts [3,4]. Moreover, several studies have shown beneficial effects for organizations [5,6]. Despite the great impact of social media on society, and even though its use in blood donor recruitment and retention has increased [7,8], the role and impact of social media have been overlooked in donor research. Maintaining a stable blood supply is essential to worldwide health systems [9,10]. With less than 5% of the eligible population donating blood [11], it is essential to retain current and recruit new donors. Social media might be a powerful tool to accomplish this. Therefore, we aim to map the social media landscape around blood donation and investigate dynamic content developments over time in order to identify key topics and momentum to inform strategic blood donor management.

Social Media Background

Social media have transformed the way people interact with organizations, enabling two-way communication in which people can act and react to organizations [12–14]. This has made social media an increasingly necessary tool for organizations to engage online and enhance their stakeholders' trust, attitudes, and commitment [15]. This is certainly true for nonprofit organizations. Attention is a scarce and essential resource for nonprofit organizations to achieve impact and reach their goals [1]. Until recently, organizations were reliant on gatekeepers (who determine which stories

* Correspondence to: Steven Ramondt, Department of Communication Science, Vrije Universiteit Amsterdam, De Boelelaan 1101, 1081 HV Amsterdam, the Netherland.

E-mail address: s.ramondt@sanquin.nl (S. Ramondt).

https://doi.org/10.1016/j.tmrv.2021.10.001

receive attention) to reach large audiences and bring notice to their cause in the media [16]. Social media have empowered nonprofits in this quest for attention: the interactive, decentralized environment of social media allows for relationship building, engaging stakeholders, and focusing attention on messages that traditional media might ignore as competition for media attention is fierce [17]. Organizations do not only benefit from the large audiences available but also from the potential to spread awareness to a diverse audience rapidly [18]. This allows nonprofits to reach new audiences that can be mobilized to take action [19]. The nonprofit sector has therefore adopted social media to aid their cause by facilitating fundraising, advocacy, community building, and other beneficial activities [19–21].

Social media have shown promise for collection agencies of substances of human origin, including blood and organ donation. For example, a campaign focused on social norms showed a rapid -though temporary- 700 percent spike in new organ donors in the US by enabling the sharing of organ donor status to friends on Facebook [4]. In the Netherlands, online recruitment and social media in particular, have become the primary focus of Dutch donor recruitment and retention [8]. Several recent studies have investigated social media's potential for stimulating blood donation behavior. For instance, a study conducted in South Africa showed a positive relationship between social media communications and blood donation awareness, personal perceptions, perceptions of family and peer influence, and donation intention [22]. Studies conducted in India [23] and Saudi Arabia [24] have shown social media's promise as a tool to spread blood donation requests. In a German survey, social media was reported to be the second most important motivation to donate and the primary motivational factor for first-time donors [25]. However, it remains unclear what characteristics and content of social media are responsible for the motivating effect.

To make optimal use of social media for recruitment and retention campaigns, understanding the environment in which they take place is crucial. Social media is a turbulent environment in which experiences and opinions are freely and abundantly shared. These experiences and opinions are often publicly visible and therefore add another dimension beyond the traditional communication attempts between organizations and their target audience [26]. People have been progressively valuing opinions of peers, online reviews, and other forms of electronic word-of-mouth, possibly impacting the reception and effectiveness of communication campaigns [27,28]. A better understanding of the dynamics of opinions about blood donation on social media will help blood collection agencies to make strategic choices regarding important dimensions of communication strategies such as timing, pacing, and content [1]. New campaigns can address timely topics and growing concerns while recognizing what donors and non-donors have to say about blood donation matters. Long-term planning based on consequences of trending and changing opinions about blood donation becomes feasible. In addition, timing of (social media) campaigns and how often messages are sent can be adjusted to the social media landscape, especially when this landscape is monitored real-time.

Blood Donation Opinions and Experiences

Insights into opinions about blood donation, donor experiences and motivations, and donation consequences are critical for targeted and tailored donor recruitment and retention. Recognition of donor research as essential for a healthy blood supply has led to multiple observational and survey studies to acquire these insights. However, the current literature is primarily focused on understanding donors and therefore lacks information about non-donors, crucial for recruitment. Moreover, these studies consist mainly of extensive surveys that are expensive, primarily in donor burden. Even though large-scale donor survey studies remain imperative to provide us with crucial donor information, we cannot rely on these studies alone to respond to sudden changes in opinions about blood donation that require a response from blood collection agencies. The rise of social media has allowed us to supplement donor research with a dynamic data source that can fill some of the gaps and limitations that currently exist. It is possible to mine social media messages to continuously gain insights into opinions about blood donation from donors as well as non-donors without study-specific response bias and without participation burden.

Social media is an unstructured data source that is big, rich, and always-on [29]. As a result, social media can provide us with new insights. For example, we recently found barriers in social media messages that have been underexposed in survey research and found peaks in new donor registrations during the COVID-19 pandemic coincident with peaks in (social) media attention [7,30]. However, the abundance of data that social media provides comes in an unstructured form (ie, short unordered textual responses). Without extensive extraction, pre-processing, and analyzing these data may not lead to improved decision making, interventions, and ultimately improved performance [31]. New opportunities in the field of computational social science allow us to do so longitudinally and in detail [29].

The Current Study

We present a methodology that blood collection agencies can use to better understand opinions of donors and non-donors about blood donation found on social media. We aim to convert the unstructured blood donation data provided by social media into a structured form, and summarize and review the content. Specifically, we apply a topic modeling methodology that allows us to utilize unsupervised machine learning to discover the latent topics in the data and how they change over time. As such, we "automatically" code our social media data for the existence of topics, allowing us to analyze large amounts of documents and text without manual coding.

With topic modeling, we take advantage of computer algorithms that use frequency distributions of words to identify latent patterns of word occurrence [32]. This provides us with clusters of co-occurring words with a shared subject matter in a collection of documents. These clusters represent sets of topics. As is the case with all latent variables, the interpretation of the variable is up to the researcher. In the case of topic modeling, this means identifying the underlying topics in clusters of words [33]. It is worth noting that this means our study is inductive, exploratory, and does not allow us to form a solid set of hypotheses. Instead, our purpose is aimed at better understanding the state of blood donation messages on social media while utilizing prior findings and theory to aid interpretation of the results. Specifically, employing social media data collected among the Dutch population, we aim to answer the following questions: (1) What topics do Dutch social media users discuss when they talk about blood donation on social media, (2) how are the discussed topics distributed on Facebook and Twitter, and (3) how does the prevalence of topics change over time?

To answer these questions, we employ a Structural Topic Model (STM) [34], an extension to the Latent Dirichlet Allocation (LDA) [35] topic modeling methodology that has been ground-breaking for text analysis and computational linguistics. STM allows for the incorporation of covariates that make it possible to outperform other approaches of topic modeling [34]. Moreover, including time and platforms as covariates provides us with valuable information related to the topic distribution, prevalence, and dynamics.

Material and Methods

Data Collection

We explored Dutch public Facebook and Twitter posts from January 1, 2012, when blood donation gained traction on Dutch social media, until 31 December, 2018. To collect our data, we used the social media tool Coosto (social media management software: www.coosto.nl) and applied a search string to their extensive social media database with public Dutch social media messages. This process to collect social media messages is similar to a literature review search. We included various (combinations of) terms related to blood donation and Sanquin in Dutch language posts only (see Appendix A, Table A1). Facebook and Twitter contain the most public messages about blood donation and are the platforms on which Sanquin (the Dutch blood collection agency) is most active. Nearly 40% of the messages about blood donation included in this study were either posts or replies by individuals on the Sanquin Facebook page or replies to or mentions by individuals of the Sanquin Twitter account. After we collected our social media messages, we imported them in R for further cleaning. To get to our final corpus of documents, we omitted messages that blood collection agencies posted. We further removed duplicates and usernames identified as expressing irrelevant information (eg, job sites and veterinarian blood banks, see Appendix A, Table A2). No additional messages were removed as the algorithm identifies the underlying topics, and topics that are not relevant to this research (eg, illegal blood transfusions, religious texts, and blood collection in Belgium) can be ignored.

Procedure

Our procedure consisted of four steps. First, we conducted data pre-processing steps to prepare our data for analyses. Second, we assessed the optimal number of topics based on statistical indicators. Third, we used two coders to decide on the final corpus, number of topics, and label of the topics. Lastly, we manually clustered the topics.

Pre-Processing

Our analyses and pre-processing were performed in R [36], and the pre-processing steps were conducted using the corpustools package. To improve results and decrease unnecessary computing time it is recommended to reduce the data in pre-processing steps. Our pre-processing techniques followed established recommendations [32,37] for topic modeling and consisted of the following steps: (1) lemmatization. While tokenization is the common procedure for English text pre-processing (only the stem of a word is left), for richly inflected languages such as Dutch, lemmatization (using a lexicon in combination with regular conjugation) is more appropriate [32], (2) removal of stop words and elimination of numbers and punctuation, (3) retention of words with minimal 3 characters, (4) additional data reduction. To filter out rare terms we only kept terms that occur minimally 20 times. Rare terms do not appear in the cluster of words that represent topics and can therefore be removed without significant consequences. In addition, we only include terms that occur in a maximum of 25% of the documents because overly common terms will not contribute to meaningful and distinguishable topics.

We took additional steps because the data consist of social media messages. Topic modeling may cause problems in extracting meaningful topics as social media messages are short and topic models need to be reasonably large to lead to optimal results [38]. To deal with this problem, multiple aggregation solutions have been suggested to increase document size, including aggregating posts by the same author over time [39], aggregating posts in a certain time frame [40], and combinations thereof [37]. However, no decisive standard for dealing with the difficulties of social media data in topic modeling has been agreed upon. We therefore tested multiple pre-processing plans for our optimal solution. We looked at a solution without aggregation, one where we aggregated messages over time (combining messages in a 1-day period across users), and a new solution based on Guo et al. (2016) [37] in which we combined only included users with more than 4 posts and combined every 4 messages of a user. We further compared these preparation steps with another step in which we applied part of speech (POS) tagging. POS-tagging indicates keeping only specific types of words (eg, verbs or nouns) in the data. Some studies recommend using only specific parts of sentences that contain the most important information to be included in the topic model, and therefore recommend POS-tagging [32,41]. We utilized UDpipe [42] to conduct POS-tagging with the corpustool package and kept nouns, adjectives, and proper nouns as we expect these types to contain the most relevant information for blood donation topics.

Our additional steps left us with 2 (POS vs no POS) x 3 (no aggregation, aggregation over time, and aggregation over posts by the same user) pre-processing plans. The inspection of the results by the coders showed that the STM with no POS and no aggregation was most optimal for these data (see Appendix A for details on these additional pre-processing steps). Only this pre-processing plan will be discussed in the main paper.

Topic Estimation

We calculated and visualized the STM utilizing the stm package. Time (in days) and social media platform (Facebook or Twitter) were included as covariates. In order to evaluate topics, we specified the number of topics (k) the model should classify. We started with 5 topics and incrementally added 5 topics to a maximum of 50 topics. The statistical indicators for the optimal number of topics were based on three criteria set by Roberts and colleagues (2019) [43]: (1) held out likelihood, which explains the overall variability in the corpus and is preferred to be as low as possible, (2) semantic coherence, which is highest when probable words in each topic frequently co-occur, and (3) exclusivity of words to the topic. We chose 2 solutions (ie, 15 and 50 topics) that had the best relationship between these three criteria and manually inspected these 2 solutions.

Topic Interpretation

The topic model algorithm can find general as well as context-specific topics with a refinement that might not be meaningful for humans, such as linguistic artifacts [33,44]. It is therefore important that the number of included topics is based on both statistical indicators as well as decisions by the researcher. We used the Stmingsights package to look at most common words, FREX words (word ranking based on semantic coherence and exclusivity), and visualizations for each topic to determine the optimal number of topics and topic labels. The topic solution with 50 topics showed the most distinct and meaningful topics. The 15-topic solution unnecessarily combined distinct topics and was missing meaningful topics. Half of the topics in the 50-topic solution were considered irrelevant (eg, religious texts and blood collection in Belgium), leaving us with 25 topics. We manually labeled these topics.

Topic Clustering

We visualized the topics in a correlation network to help the researchers cluster the topics and enhance understanding of the

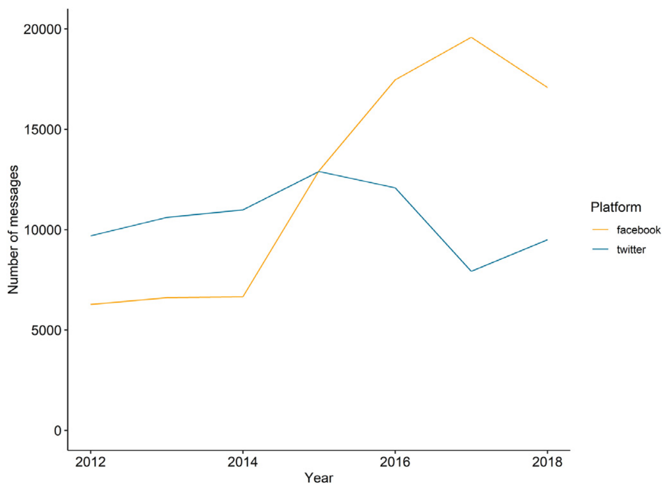

Fig. 1. Messages over time by platform.

Table 1
STM results for blood donation on social media

Cluster	Topic label	Prop.
1) Announcements		0.136
	Donation announcement: image motivation	0.070
	Donation announcement: hashtag	0.036
	Donation announcement: regular	0.030
2) Non-donation reasons and deferral		0.161
	Reasons for non-donation	0.045
	Deferral	0.041
	Deferral: transfusion	0.041
	Fear of needles	0.015
	Negative experiences	0.011
	Donor selection	0.008
3) Donation process		0.127
	Plasma donation	0.046
	Invitation procedure	0.038
	Donation snack	0.024
	Appointment & location	0.019
4) Positives of donation		0.076
	Donor identity	0.033
	Matching (blood type)	0.033
	Donor appreciation	0.019
	Patient story	0.015
	Attention for donation	0.009
5) Controversial policies		0.066
	MSM: policy decision	0.032
	MSM: discussion	0.027
	Remuneration	0.007
6) Campaigns and registration		0.083
	World blood donor day	0.024
	Blood saves lives	0.023
	Donor registration (not only blood)	0.022
	Missing type campaign	0.014

topics and their relationship with each other. Then, we manually grouped the 25 labeled topics into 6 clusters.

Results

Our corpus contained a total of 160.313 social media messages. Figure 1 displays the total number of messages per year for each platform. It consisted of a nearly even social media platform split, with 86.616 (54%) messages on Facebook and 73.697 (46%) messages on Twitter. Figure 2 shows the 25 labeled topics in a correlation network.

Table 1 shows the estimated topic solution, our manually assigned English labels to the 25 identified topics, the proportion topics made up in our documents, and our manually assigned 6 clusters.

Cluster 1: Announcements, consists of 3 announcement topics. In the “Donation announcement: image motivation” topic, people emphasize the good deed of their donation. “Donation announcement: hashtag” consists of donation announcements with many hashtags and is the result of the algorithm separating Twitter posts with announcement messages as a distinct topic. “Donation announcement: regular” consists of donation announcement messages that have no distinctive features.

Cluster 2: Non-donation reasons and deferral, is made up of three deferral-related topics. “Deferral: transfusion” focuses on deferral due to (having received a) transfusion, “Donor selection” consists of discussions around donor eligibility, and a general “Deferral” topic, including all kinds of deferral reasons. Furthermore, the cluster includes “Reasons for non-donation” in which people explain why they cannot or do not donate. “Fear of needles” includes both messages about fear as well as vasovagal reactions as a result of blood donation, and “Negative experiences” which is a topic with a variety of negative experiences (eg, experiences with blood bank staff) that can lead to non-donation.

Cluster 3: Donation process, contains “Invitation procedure” with primarily questions and remarks about donation invitations, “Plasma donation” a broad plasma donation topic that also entails the different rules and procedures compared to whole blood donation. In addition, it includes “Appointment and location” containing questions and messages about appointments and locations, and “Donation snack” the sharing of the post-donation snacks and drinks.

Cluster 4: Positives of donation, is composed of “Donor identity” consisting of messages in which people identify themselves as donor, and “Matching” about blood type and compatibility. In addition, it includes “Donor appreciation” thanking donors, “Patient story” made up of messages about people benefiting from blood donation, and “Attention for donation” consisting of messages that ask for attention for blood donation.

Cluster 5: Controversial policies comprise controversial topics and policies that often lead to negative attention. These topics are less connected and have a weaker relationship with other topics, ensuring that they are on the outside of the network. This cluster includes 2 men who have sex with men (MSM) topics; a “MSM: policy decision” focused topic that includes legal and political decisions and news, and a broader and sometimes more personal “MSM: discussion” topic. The cluster further includes “Remuneration” in which both remuneration for blood (products) and upper management salaries are discussed and complained about.

Cluster 6: Campaigns and registration, consists of the Sanquin campaigns “Blood saves lives” and “Missing type”, and the broader donor topics “World blood donor day” and “Donor registration”.

Cluster 1 *Donation announcements* and cluster 2 *Non-donation reasons and deferral* are the clusters with the most discussed topics. The five topics containing most messages are “Donation announcement: image motivation”, “Plasma donation”, “Reasons for non-donation”, “Deferral”, and “Deferral: transfusion”.

Platform Prevalence

Figure 3 shows the topic prevalence contrast between Facebook and Twitter. We see that “MSM: policy decision” and “Announcement: hashtag” have a particularly strong prevalence on Twitter. Other topics with a more substantial prevalence on Twitter are the “Blood saves lives” and the “Missing type” campaigns. The “Deferral: transfusion” and “Donor identity” topics can be found mainly on Facebook. Overall, we see topics belonging to the clusters *Campaigns and registration* and *Controversial policies* primarily on Twitter. On Facebook, the *Donation process* cluster with topics such as “Invitation procedure” and “Donation snack” has a strong presence. Similarly, the *Positives of donation* cluster, with topics such

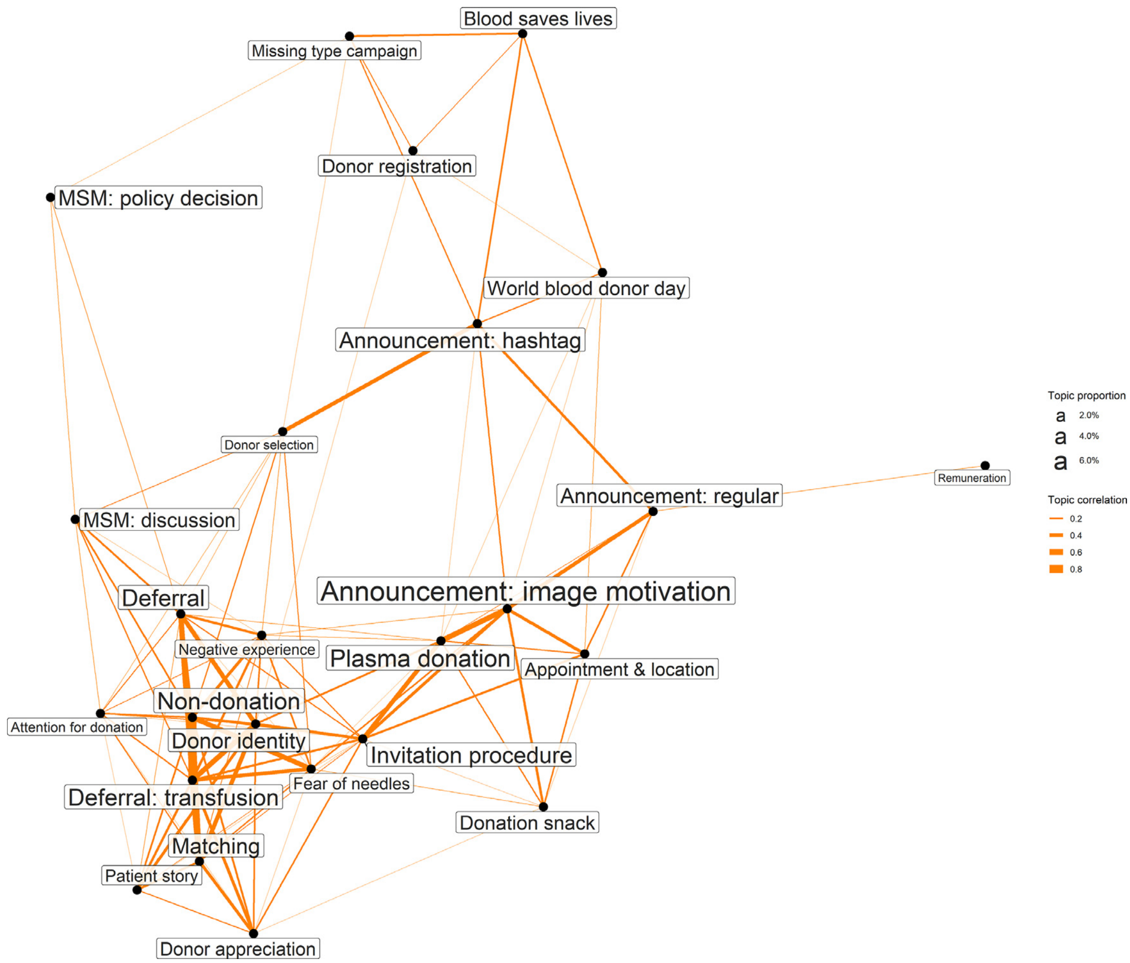

Fig. 2. Correlation network of topics.

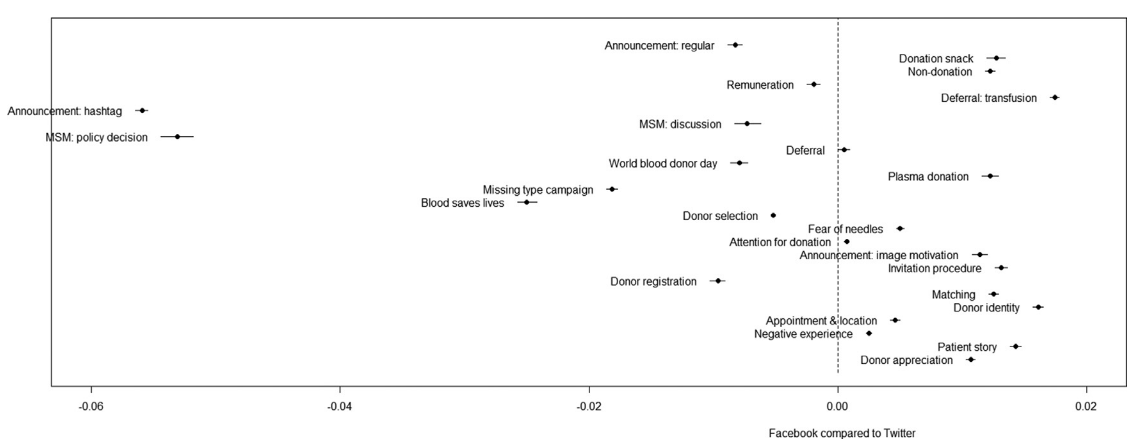

Fig. 3. Topic prevalence contrast between Facebook and Twitter. The x-axis represents prevalence with everything >0.00 indicating the topic is more prevalent on Facebook, and everything <0.00 indicating the topic is more prevalent on Twitter. Topics are distributed along the y-axis.

as "Donor identity" and "Patient story", and *Non-donation reasons and deferral*, with the deferral topics, are more prevalent on Facebook. The non-hashtag *Donation announcements* topics are divided more evenly on both platforms.

Topics Over Time

Figure 4 shows topic prevalence over time, ranging from January 1, 2012, until 31 December 2018, focusing on the topics that show an upward or downward trend. Most topics remained relatively stable throughout the period of interest with the exception of specific campaigns that peaked during their existence. However, some of the topics show upwards or downwards trends. There is a downward trend of two out of three donation announcement topics, although there is some recovery during 2018. *Positives of donation* messages are growing, with the "Donor identity", "Matching", and "Donor appreciation" topics showing an increase that is particularly strong at the end of 2018. Moreover, we see the donation snack increase, with the "pink glazed cookie" reaching mythi-

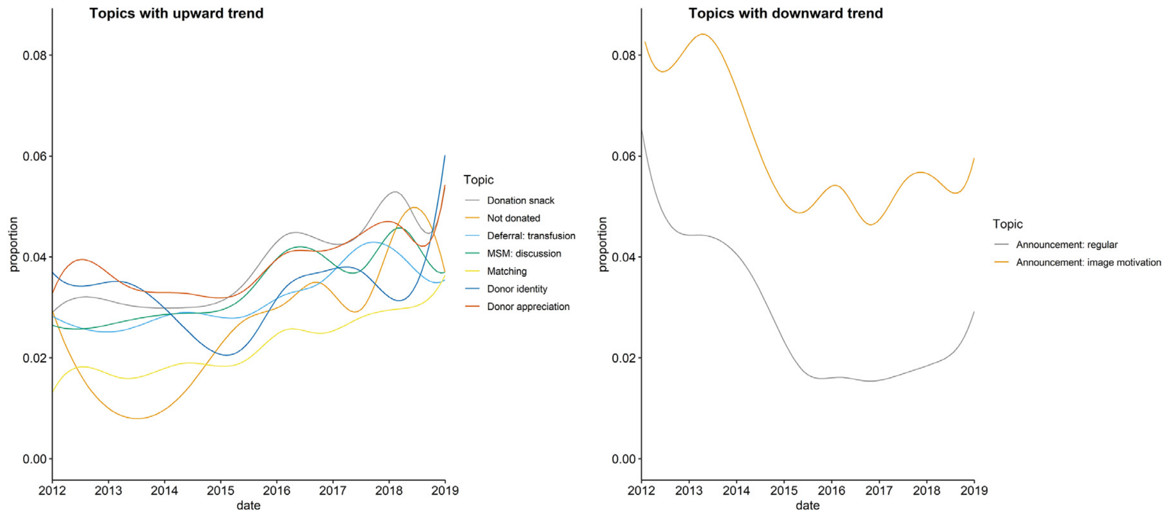

Fig. 4. Topic prevalence over time.

cal proportions in the Dutch blood donor community. There is also an increase in discussion about the MSM exclusion policy that has been increasingly under pressure. In addition, there is an upward trend of expressing why people cannot donate as we can see with the "Deferral: transfusion", and the "Non-donation" topics.

Discussion

In this paper, we showcase an inductive computational method to make sense of vast amounts of dynamic unstructured blood donation text data that exists on social media. We applied topic modeling on 7 years of Facebook and Twitter data to examine topics discussed on social media about blood donation by the general public. In our data with a particularly homogenous topic (ie, blood donation), STM without POS-tagging and no aggregation preprocessing performed optimally. This suggests that adding covariates to the model can be enough to handle potential issues that arise from utilizing social media in topic modeling. The STM further allowed us to establish topics and measure both platform prevalence and change over time that would otherwise be hard to identify.

People have embraced social media to talk about blood donation. We found 25 topics clustered in 6 distinct clusters. Positive social norms around blood donation are present on social media as people actively share that they have donated and why they can or could not. In this context, the decrease in donation announcements over time might be problematic. Beneficial social norms around blood donation might be jeopardized if donors no longer share their donation experiences online. Image motivation announcements in which donors emphasize the good deed of their donation can be perceived as bragging and have reputational costs. These costs can outweigh benefits and lead to adverse consequences [45,46]. Reassuring is the strong presence of topics that emphasize the positives of donation and the active sharing of campaigns designed by Sanquin to recruit new donors. The upward trend of topics such as donor identity, appreciation, and even donor snack is also encouraging. While banal sounding, donation snack messages are often used as an announcement statement and even appear to be part of donor identity for some donors. Trending topics such as donor identity and donor snack seem to replace some of the drop in donation announcements placed online. By making donation announcements in a manner that does not emphasize the good deed, for example by referring to donations snacks, reputational cost might be diminished.

Yet not all topics are positive and there is also an active discussion of topics related to MSM and remuneration, aimed at changing policy. For example, MSM policy is increasingly discussed online. As seen in our previous study [7] controversial policy topics such as MSM and remuneration take place on Twitter. Recruitment campaigns were also found more on Twitter. It is plausible that platform characteristics are responsible for the difference in topic prevalence on Facebook and Twitter. Twitter is an asymmetrical social media network, that does not require both parties to establish a formal link. Facebook on the other hand, is a symmetrical network in which both parties need to agree to establish a relationship [1]. Reaching people outside of your network with Twitter is therefore simpler and this makes it an effective tool to reach large and new audiences for donor recruitment and policy change. The donation procedures cluster was found more on Facebook. It is possible to ask questions on the Sanquin Facebook page, and Sanquin stimulates the use of this feature. It is therefore not surprising that questions and remarks about procedures are more prevalent on this platform. It is plausible that Sanquin has influenced topic emergence. Many of the social media posts were found on the accounts of, or in reply to Sanquin and the majority of topics are in line with the 5 social media communication pillars of Sanquin (ie, Donor Community, Patient Better Life, Sanquin Inside, Blood, and Plasma).

The difference in platform characteristics also has implications for blood collection agencies. The symmetrical aspect of Facebook makes it more likely to consist of closer relationships. Given the social contagious effect that important members of donors' social network have on blood donation behavior [47], Facebook is a promising network for community building and donor-recruits-donor campaigns. Twitter as an asymmetrical network might be better used for reaching large and new audiences. The insights from this study have implications for practice beyond platform-specific recommendations. Big data are an increasingly important tool for organizations to understand needs [48]. Blood collection agencies and other organizations should use the available data online to identify topics and trends and adjust their communication strategy when necessary. The identification of clusters of and relationships between topics can be utilized for integrated communication efforts. Our study shows a clear social media platform contrast for many topics and prime topics that can be employed to build a stronger donor community. For example, it could prove effective for Sanquin to stimulate the active sharing of donations on social media making people in the surroundings of donors aware of blood donation. This should occur on Facebook, especially given the decreasing announcements on this platform and the increased likelihood of reaching key demographics such as family and friends. It would also do well to keep building community on social media, utilizing topics such as donation snacks for this purpose, and aim to maintain the growing trend of people identifying themselves as donor.

In addition to the benefits gained by utilizing an unsupervised learning algorithm on a novel data source, our study also has specific limitations. Social media research is not immune to many of the methodological difficulties of survey research such as participation- and social-desirability bias [29,49,50]. Moreover, it is difficult to differentiate between donors and non-donors. Social media data should therefore not be the only source of opinions about blood donation and the discovered topics should not be regarded as the final reflection of blood donation topics that can be found on social media. As with all latent variables, the themes of the topics are subject to researcher interpretability. In addition, the final decision on the number of topics lies with the researcher. These results can contradict statistical indicators, and can therefore appear arbitrary [33]. Moreover, by deciding on our number of topics we determined the granularity of these topics. A higher topic solution might result in the split of a topic into other specific and meaningful topics. Finally, it is important to note that our analyses only cover Dutch public Facebook and Twitter posts, limiting our generalizability. It is plausible that Sanquin had a large influence on the topic emergence and certain topics such as the discussion around the remuneration (of upper management) might be unique to the Netherlands [7]." We therefore recommend that (other) blood collection agencies regularly conduct these analyses themselves to understand how donors and non-donors talk about blood donation on social media and how the narratives change over time. Sharing these results would allow us to identify both universal and culture-specific blood donation themes on social media. Excellent tutorials on conducting STM analyses such as by Roberts and colleagues [34,43] are available. In addition, we recommend that researchers and blood collection agencies make use of the preprocess steps and decisions we have provided that led to the most optimal results for our data.

Conclusion

We demonstrated a means of analyzing dynamic social media data about blood donation utilizing an inductive computational method. To our knowledge, this is the first study to describe what gets discussed on social media about blood donation. As such, it fills a gap in donor research and contributes to the knowledge about public opinion regarding blood donation. Moreover, it aids blood collection agencies in making strategic choices about utilizing social media to promote blood donation by providing a better understanding of the blood donation narrative on social media.

Conflict of Interest

None.

Funding

This work was funded by Sanquin Blood Supply (PPOC 18/13). EMM's contribution has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement 802227).

CRediT authorship contribution statement

Steven Ramondt: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Peter Kerkhof:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing – review & editing. **Eva-Maria Merz:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – review & editing.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tmrv.2021.10.001.

CRediT authorship contribution statement

Steven Ramondt: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **Peter Kerkhof:** Conceptualization, Data curation, Funding acquisition, Supervision, Writing – review & editing. **Eva-Maria Merz:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – review & editing.

References

[1] Guo C, Saxton GD. The quest for attention: nonprofit advocacy in a social media age. Stanford, California: Stanford University Press; 2020.
[2] Pew Research CenterDemographics of social media users and adoption in the United States. Soc Media Fact Sheet 2019. https://www.pewresearch.org/internet/fact-sheet/social-media/ (accessed March 29, 2021).
[3] Bail CA. Cultural carrying capacity: organ donation advocacy, discursive framing, and social media engagement. Soc Sci Med 2016;165:280–8. doi:10.1016/j.socscimed.2016.01.049.
[4] Cameron AM, Massie AB, Alexander CE, Stewart B, Montgomery RA, Benavides NR, et al. Social media and organ donor registration: the Facebook Effect. Am J Transplant 2013;13:2059–65. doi:10.1111/ajt.12312.
[5] Kim AJ, Ko E. Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. J Bus Res 2012;65:1480–6. doi:10.1016/j.jbusres.2011.10.014.
[6] Beukeboom CJ, Kerkhof P, de Vries M. Does a virtual like cause actual liking? How following a brand's Facebook updates enhances brand evaluations and purchase intention. J Interact Mark 2015;32:26–36. doi:10.1016/j.intmar.2015.09.003.
[7] Ramondt S, Zijlstra M, Kerkhof P, Merz E-M. Barriers to blood donation on social media: an analysis of Facebook and Twitter posts. Transfusion (Paris) 2020;60:2294–306. doi:10.1111/trf.15998.
[8] van Ark T. Ministerieel Plan Bloedvoorziening 2021-2023, 2021. https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2021/02/25/kamerbrief-over-ministerieel-plan-bloedvoorziening-2021-2023/kamerbrief-over-ministerieel-plan-bloedvoorziening-2021-2023.pdf(accessed May 19, 2021).
[9] Healy K. Embedded Altruism: blood collection regimes and the European Union's donor population. AJS 2000;105:1633–57.
[10] Slonim R, Wang C, Garbarino E. The market for blood. J Econ Perspect 2014;28:177–96. doi:10.1257/jep.28.2.177.
[11] To L, Dunnington T, Thomas C, Love K, McCullough J, Riley W. The United States' potential blood donor pool: updating the prevalence of donor-exclusion factors on the pool of potential donors. Transfusion (Paris) 2020;60:206–15. doi:10.1111/trf.15573.
[12] Hanna R, Rohm A, Crittenden VL. We're all connected: the power of the social media ecosystem. Bus Horiz 2011;54:265–73. doi:10.1016/j.bushor.2011.01.007.
[13] Dijkmans C, Kerkhof P, Beukeboom CJ. A stage to engage: social media use and corporate reputation. Tour Manag 2015;47:58–67. doi:10.1016/j.tourman.2014.09.005.
[14] Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Bus Horiz 2010;53:59–68. doi:10.1016/j.bushor.2009.09.003.
[15] van Noort G, Willemsen LM. online damage control: the effects of proactive versus reactive webcare interventions in consumer-generated and brand-generated platforms. J Interact Mark 2012;26:131–40. doi:10.1016/j.intmar.2011.07.001.
[16] Bail CA. Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. Proc Natl Acad Sci 2016;113:11823–8. doi:10.1073/pnas.1607151113.
[17] Lovejoy K, Saxton GD. Information, community, and action: how nonprofit organizations use social media. J Comput-Mediat Commun 2012;17:337–53. doi:10.1111/j.1083-6101.2012.01576.x.
[18] Bakshy E, Messing S, Adamic LA. Exposure to ideologically diverse news and opinion on Facebook. Science 2015;348:1130–2. doi:10.1126/science.aaa1160.
[19] Guo C, Saxton GD. Tweeting social change: how social media are changing nonprofit advocacy. Nonprofit Volunt Sect Q 2014;43:57–79. doi:10.1177/0899764012471585.
[20] Campbell DA, Lambright KT, Wells CJ. Looking for friends, fans, and followers? social media use in public and nonprofit human services. Public Adm Rev 2014;74:655–63. doi:10.1111/puar.12261.

[21] Young JA. Facebook, twitter, and blogs: the adoption and utilization of social media in nonprofit human service organizations. Hum Serv Organ Manag Leadersh Gov 2017;41:44–57. doi:10.1080/23303131.2016.1192574.
[22] Duh HI, Dabula N. Millennials' socio-psychology and blood donation intention developed from social media communications: a survey of university students. Telemat Inform 2021;58:101534. doi:10.1016/j.tele.2020.101534.
[23] Abbasi RA, Maqbool O, Mushtaq M, Aljohani NR, Daud A, Alowibdi JS, et al. Saving lives using social media: analysis of the role of twitter for personal blood donation requests and dissemination. Telemat Inform 2018;35:892–912. doi:10.1016/j.tele.2017.01.010.
[24] Alanzi T, Alsaeed B. Use of social media in the blood donation process in Saudi Arabia. J Blood Med 2019;10:417–23. doi:10.2147/JBM.S217950.
[25] Sümnig A, Feig M, Greinacher A, Thiele T. The role of social media for blood donor motivation and recruitment. Transfusion (Paris) 2018;58:2257–9. doi:10.1111/trf.14823.
[26] Kerkhof P, Dijkmans C. Webcare. Oxford bibliographies in communication. New York: Oxford University press; 2019.
[27] Park DH, Lee J, Han I. The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: the Moderating Role of Involvement. Int J Electron Commer 2007;11:125–48. doi:10.2753/JEC1086-4415110405.
[28] Utz S, Kerkhof P, van den Bos J. Consumers rule: how consumer reviews influence perceived trustworthiness of online stores. Electron Commer Res Appl 2012;11:49–58. doi:10.1016/j.elerap.2011.07.010.
[29] Salganik MJ. Bit by bit: social research in the digital age. Princeton: Princeton University Press; 2017.
[30] Spekman MLC, Ramondt S, Quee FA, Prinsze FJ, Veld EMJH in 't, Hurk K van den, et al. New blood donors in times of crisis: increased donation willingness, particularly among people at high risk for attracting SARS-CoV-2. Transfusion 2021;61:1822–9. doi:10.1111/trf.16334.
[31] Xu Z, Frankwick GL, Ramirez E. Effects of big data analytics and traditional marketing analytics on new product success: a knowledge fusion perspective. J Bus Res 2016;69:1562–6. doi:10.1016/j.jbusres.2015.10.017.
[32] Jacobi C, Atteveldt W van, Welbers K. Quantitative analysis of large amounts of journalistic texts using topic modelling. Digit Journal 2016;4:89–106. doi:10.1080/21670811.2015.1093271.
[33] Chang J, Boyd-Graber J, Gerrish S, Wang C, Blei DM. Reading tea leaves: how humans interpret topic models. In: Proc. 22nd Int. Conf. Neural Inf. Process. Syst., Red Hook, NY, USA: Curran Associates Inc.; 2009. p. 288–96.
[34] Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural topic models for open-ended survey responses. Am J Polit Sci 2014;58:1064–82. doi:10.1111/ajps.12103.
[35] Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. J Mach Learn Res 2003;3:993–1022.
[36] R Core Team R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
[37] Guo L, Vargo CJ, Pan Z, Ding W, Ishwar P. Big social data analytics in journalism and mass communication: comparing dictionary-based text analysis and unsupervised topic modeling. Journal Mass Commun Q 2016;93:332–59. doi:10.1177/1077699016639231.
[38] Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the limiting factors of topic modeling via posterior contraction analysis. In: Proc. 31st Int. Conf. Int. Conf. Mach. Learn. - Vol. 32, Beijing, China: JMLR.org; 2014. p. 190–8.
[39] Hong L, Davison BD. Empirical study of topic modeling in Twitter. In: Proc. First Workshop Soc. Media Anal., New York, NY, USA: Association for Computing Machinery; 2010. p. 80–8. doi:10.1145/1964858.1964870.
[40] Zhao WX, Jiang J, Weng J, He J, Lim E-P, Yan H, et al. Comparing Twitter and Traditional Media Using Topic Models. In: Clough P, Foley C, Gurrin C, Jones GJF, Kraaij W, Lee H, et al., editors. Advances in information retrieval. Berlin, Heidelberg: Springer; 2011. p. 338–49. doi:10.1007/978-3-642-20161-5_34.
[41] Korfiatis N, Stamolampros P, Kourouthanassis P, Sagiadinos V. Measuring service quality from unstructured data: a topic modeling application on airline passengers' online reviews. Expert Syst Appl 2019;116:472–86. doi:10.1016/j.eswa.2018.09.037.
[42] Straka M, Straková J. Tokenizing, P.O.S. Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. In: Proceedings of the CoNLL 2017 Shared Task. Vancouver, Canada: Association for Computational Linguistics; 2017. p. 88–99. https://www.aclweb.org/anthology/K17-3009. Accessed March 23, 2021. doi:10.18653/v1/K17-3009.
[43] Roberts ME, Stewart BM, Tingley D. stm: an R package for structural topic models. J Stat Softw 2019;91:1–40. doi:10.18637/jss.v091.i02.
[44] Mimno D, Blei D. Bayesian checking for topic models. In: Proc. Conf. Empir. Methods Nat. Lang. Process., USA: Association for Computational Linguistics; 2011. p. 227–37.
[45] Graf C., Merz E.-.M., Suanet B., Wiepking P. Social Norms Offer Explanation for Inconsistent Effects of Incentives on Prosocial Behavior. ArXiv210413652 Econ Q-Fin 2021.
[46] Scopelliti I, Loewenstein G, Vosgerau J. You Call It "Self-Exuberance"; I Call It "Bragging": miscalibrated Predictions of Emotional Responses to Self-Promotion. Psychol Sci 2015;26:903–14. doi:10.1177/0956797615573516.
[47] Schröder JM, Merz E-M, Suanet B, Wiepking P. The social contagion of blood donations: sustaining repeated prosocial behaviour. SocArXiv 2021. doi:10.31235/osf.io/kjcwh.
[48] CPI.Can big data protect a firm from competition? Compet Policy Int 2017. Accessed March 29, 2021 https://www.competitionpolicyinternational.com/can-big-data-protect-a-firm-from-competition/ .
[49] Schober MF, Pasek J, Guggenheim L, Lampe C, Conrad FG. Social Media Analyses for Social Measurement. Public Opin Q 2016;80:180–211. doi:10.1093/poq/nfv048.
[50] Heiervang E, Goodman R. Advantages and limitations of web-based surveys: evidence from a child mental health survey. Soc Psychiatry Psychiatr Epidemiol 2011;46:69–76. doi:10.1007/s00127-009-0171-9.