A Tipping Point? Heightened self-disclosure during the Coronavirus pandemic

Anna Squicciarini, Sarah Rajtmajer, Prasanna Umar, Taylor Blose
College of Information Sciences and Technology,
Pennsylvania State University,
University Park 16802
{acs20, smr48, pxu3, thb5018}@psu.edu

2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI) | 978-1-7281-4144-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/COGMI50398.2020.00040

***Abstract*—The COVID-19 crisis has raised numerous concerns amongst the privacy community, as research indicates emerging privacy risks. Here we frame those risks and focus, in particular, on what we have observed to be increased rate of self-disclosure on online social media. That is, individuals are sharing more personal information online in what appears to be an effort to stay connected with others during isolation and cope with additional stress. We outline a research agenda to further explore this finding and highlight the potential for this crisis to serve as a tipping point for self-disclosure norms more generally.**

I. Introduction

As we navigate the COVID-19 crisis and its broad impacts, we are understandably focused on what appear to be our most acute concerns – flatten the curve, save the economy, and develop a vaccine. Daily, we are reminded of the massive toll in lives and livelihoods that the pandemic has exacted worldwide. At the time of writing, more than 10 million cases of COVID-19 have been confirmed, and more than 500,000 lives have been lost. Economically, we are facing the largest global recession in decades and a historic contraction of per capita income.

But as the days and weeks push forward we are also beginning to understand more subtle impacts across communities. Many of these are consequent to the physical distancing and quarantine measures which have been put in place to slow the spread of disease. Crisis, coupled with isolation, has raised important concerns about mental health. An April 2020 poll of US adults [27] reported that 56% of respondents had experienced a negative impact on their mental health associated with worry and stress over the coronavirus. Text messages to the US government's Substance Abuse and Mental Health Services Administration's (SAMHSA) Disaster Distress Helpline increased nearly 1000 percent [8]. Data suggests a significant increase in cases of domestic violence [3], particularly among marginalized populations.

Unsurprisingly, there has also been an unprecedented surge in online activity [23, 16, 14]. Much of this activity has extended beyond typical Internet surfing and video streaming, as people continue to find ways to leverage online resources to stay connected with one another, personally and professionally [12, 25, 21, 5]. The last few months have seen a highly publicized rapid growth of videoconferencing companies and surging social media usage. Facebook has reported increases virus [33]. Twitter also reported an increase in daily active users by 23% from this quarter last year [20]. That is to say, we are living out this crisis online.

The benefits of this connectedness are countless. Internet technologies have allowed many to continue to work through this period, and children to continue schooling. Telehealth has expanded access to essential health services. Contact tracing apps and digital surveillance measures offer hope for virus tracking and containment. We have suggested that social media provides an avenue of "collective coping" wherein users can solicit and receive emotional, informational and instrumental support [24]. Online sharing also serves therapeutic functions [13] and enables sense-making in crises [24].

These benefits, however, raise concerns as well. Our group, amongst others, has suggested [4] that social distancing has increased individual privacy risks in online interactions. We have observed that social media users appear to be sharing more personal information on public platforms in recent months, and we note that this information is vulnerable to collection and misuse.

Our work in this area subscribes to the broader literature on so-called *self-disclosure*. Self-disclosure as such is defined as the voluntary revealing of personal information , e.g. motives, desires, feelings, thoughts, experiences, in exchange for social reward, i.e., engagement, support, legitimacy. Consistent with this definition, we believe that the observed social convergence and the corresponding rise in self-disclosure in online public platforms during this pandemic serve various, important functional purposes for users during this difficult time. But, simultaneously, there are meaningful risks associated with voluntarily shared personal information that may not always be apparent. Through this line of research, our group seeks to examine self-disclosure in online public platforms during the pandemic and understand associated privacy implications and concerns.

II. Understanding Privacy Risk: A Research Agenda

We identify several acute privacy concerns that are aggravated by the pandemic.

1) Contact Tracing: Privacy concerns related to contact tracing protocols have now (rightfully) gained widespread ntion. Inherently, the contact tracing process involves a

DOI 10.1109/CogMI50398.2020.00040

fine-tuned accounting of one's location history, relationships and health information – each sensitive information in its own right. When considered together, in the context of a global crisis, the listing of potential harms lengthens. A number of nonprofits and human rights organizations have raised concerns [10] about the potential for these tools to exacerbate domestic abuse, intensify stigmatization, and expose vulnerable populations. With a few notable exceptions [32], most countries are following protocols that rely on voluntary reporting [34]. In the US, a unified nationwide contact tracing protocol has not emerged, but rather exposure notification efforts are state-directed, with contact tracers and contract tracing apps apps sponsored by the state-level authorities.

Privacy advocates stress a few key requirements for contact tracing protocols. The first is transparency. Users must understand exactly what data will be collected, how it will be used, and who will have access it. The second is decentralization. That is, an individual's data should be stored only on their personal device and not on a central server. This mitigates risk of data breach or misuse at the central authority. Third, data should be destroyed after a preestablished period of time. This requirement acts in service to the first, helping to ensure data is not later used in unforeseen ways.

Ultimately, the success or failure of any opt-in contact tracing technology will depend on sufficient participation (by some estimates more than 50% of the population [34]), so that user trust will be paramount and privacy will take center stage. It is yet too early to assess whether contact tracing has been or will be successful, and the impact on users' privacy.

2) Health Data: As the pandemic unfolds, health-related data has been a key source of study for researchers, organizations and customers. Researchers have been racing to study short and long term impact of the Covid virus on people's health. Medical records, specimens and biological samples are being collected, shared and analysed in various forms, leaving privacy considerations behind.

At the individual's level, temperature and vital signs are constantly collected, and maintained with no underlying policy on data management and sharing. New forms of medicines have emerged, with telemedicine and remote services applied quickly. Individuals and medical units have resorted to HIPPAA framework for regulating and protecting the privacy health data, but it appears to be falling short in many ways, especially in light of new forms of medicine and research that the pandemic has brought to life [6].

3) Social Isolation and Self-Disclosure: Self-disclosure through social media carries routine risks even on the most ordinary days, leaving users exposed to discrimination, harassment and bullying, cyberfraud and other crimes. High-profile examples of collection and misuse of self-disclosed personal information (e.g., Cambridge Analytica scandal) forewarn the dangers of user targeting and manipulation. We argue that heightened rates of self-disclosure, coupled with increased anxiety and stress, may leave users particularly vulnerable to these harms. For one, we know that state-backed actors are already leveraging social media to spread disinformation related to the pandemic [36]. It is difficult to measure the impact of these campaigns. We can measure engagement spread, but the extent to which malicious posts meaningfully impact hearts and minds remains an open question. What is relatively certain, however, is that the efficacy of these efforts depends on accurate targeting, an art in turn facilitated by user sharing.

A. Contexts and Types of Self-Disclosure

In order to understand self-disclosure in social media during disasters (here, pandemic), we must first examine the types of *social convergence* observed. Depending on the socially convergent behaviors and individuals' motivations, sharing behaviors arise. Hughes and colleagues [19] outlined five types of socially convergent online behavior during crises: helping, being anxious, returning, supporting, exploiting, and being curious. These types of behaviors are reflected in the content analyses of social media (Twitter) [7, 9]. Of interest are the behaviors of helping, being anxious, supporting, exploiting and being curious. Helping refers to the behavior of contributing in terms of information on personal safety and relief sources. Users converging online also display a behavior of being anxious, meaning looking for information on safety and well-being of loved ones and others. Supporting behavior entails expressing gratitude to responders and giving support to victims. Users also use social media to mourn the loss of life. Exploiting parties compromise privacy of users, promote rumors and market their interests. Curious users lurk online to learn more about the situation. Aligning with these behaviors and motivations to engage in online activity, a recent study [35] points out to five specific purposes for Tweets during a disaster: seeking and giving information; media sharing; helping and fundraising; sharing direct experiences and narratives; and discussing and reacting. Further, content examination in different disasters [18, 24] including health crises [7] add validity to these motivations.

Considering the motivations behind online engagements (sharing) in relation to a disaster and the content types, self-disclosure in disaster-related contexts can be broadly categorized into *informational and emotional* types. Informational disclosure contain factual information about the user and emotional disclosure reveals the emotional state of the user. Such categorization is also in line with the broader self-disclosure literature [11, 37, 22]. Among five functions listed by [35], we believe informational disclosure can serve many functions ranging mostly from maintaining situational awareness, sharing personal experiences and adding support to one's positions in discussions.

Emotional disclosure can be associated mostly with reacting during the crises with range of emotions. Recent studies have pointed out to the rise of negative emotions during the pandemic [29] and linked the types of emotions as reactions to pandemic characteristics [15]. We will examine instances of informational and emotional disclosure in user contributed content/post during corona virus pandemic in light of the functional purpose sought by the disclosures. We will leverage contextual cues (e.g. prior comments, outer links, URLs, tone of messages) to define a possible purpose, and use it to better understand the post's role in the conversation. Additionally, we will examine the possible measures of the social rewards sought by users through their disclosures. For instance, in Twitter, counters of favoritism and number of replies/retweets can be proxy measures of rewards while in Reddit, “karma” could be a measure. We believe self-disclosure in replies could be a measure of social rewards. Previous studies on self-disclosing behavior posit that self-disclosure begets reciprocal self-disclosure which increases trust and builds relationships [2]. It has also been observed that support providers use self-disclosure in response to self-disclosure in support seeking posts [30, 31]. Hence, we seek to examine *reciprocity* in self-disclosure as a measure of social reward.

III. Ongoing Work

We are in the process of collecting two datasets: one from the Reddit platform and one from Twitter. Traces will cover data from Feb. 13, 2020 through June 2020 (at least). Our data collection strategy is as follows:

- *User-centric data.* Our datasets include user-led conversations discussing topics and personal experiences surrounding this health crisis. We will exclude authoritative sources (e.g., Twitter accounts managed by public entities). Conversations will be collected in their entirety (e.g., Tweets will include replies), along with any type of user feedback supported by the platform (e.g. retweets, likes, points etc). User ids, timestamps, and other key variables will be stored. Online social metrics, e.g., number of friends or followers/following will also be included.
- *Relevance.* We sample our conversational data by carefully selected keywords, hashtags and trending topics (e.g., community spread, social distancing, social isolation, #Istayhome). For the Reddit platform, we will focus on the `OffmyChest` subreddit, a community within Reddit that encourages users to share personal experiences and ask for advice. PIs and graduate students, along with social science colleagues in the College of Information Science and Technology, will help confirm the relevance of the keywords and search terms. This is needed to filter out spam or irrelevant comments/text.
- *Manual labeling.* To enable machine-based detection, we have manually labeled a large sample of our proposed Reddit and Twitter datasets (about 10,000 comments each). We label the following four categories of SD: “No Self-disclosure”, “Emotional Self-Disclosure” (”I feel sad about all the sick people in Italy #coronavirus”), “Informational Self-Disclosure” (”I am 25 years old and I tested positive for Covid-19”), and “Cannot decide”. This scheme follows the recent AAAI labeled dataset of the AffCon 2020 challenge [1], where baseline instructions for text labeling of SD were given. We will use Mechanical Turk to manually label data at different levels of granularity, i.e., sentences, comments, or conversations.

Machine-based detection of Self-Disclosure

We have begun developing tools to detect and label SD in online text. In a proof-of-concept study (60k users' comments), we tagged 9 categories of SD (e.g., race, locati relationships, interests, etc.) and studied relationships between the various categories of SD, the anonymity/identifiability of users, and topics of discussion [38].We anchored our approach in an opinion extraction technique [26], wherein the opinion, opinion-holder and the opinion topic are extracted. To detect objective language pertaining to SD, we leveraged both the semantic and the syntactic resources in a sentence. Overall, detection and categorization consisted of four phases: (1) dictionary construction, (2) subject-verb-object triplet detection, (3) named entity recognition and (4) rule based matching.

We will customize advanced supervised models for high accuracy and fine-grained labeling of SD. In particular, given the unique emotional challenges presented by our application domain, we will focus on developing robust approaches to separate *emotional* from *informational* disclosure. We have started investigating classification of text into two classes (SD/NSD) using Bidirectional Encoder Representations from Transformers (BERT). BERT models are suitable, as they tend to generalize well for a diverse set of NLP tasks [17]. Further, sentence-level representations obtained from the BERT model can also be fine-tuned with unlabeled datasets to better generalize on untrained data. Early results indicate that with BERT alone, on a skewed dataset (ration 2:5) of 12,860 sentences, we achieve an F1-score of 54.2. We will use this model to fine-tune word representations and for classification using sentence representations.

We will develop a set of neural networks (NN) leveraging attention mechanisms and embeddings, customized for our detection problem. CNNs (without many layers) are promising because they are inherently *parsimonious*, in that there is weight sharing. Therefore, they are more suitable for applications like ours, where we expect a limited amount of available training data. The scarcity of data labeled for SD does not support the use of other models (e.g., Deep Neural Networks) with many layers and a large number of free parameters to learn. A CNN structure will process users' comments independently while accounting for possible conversational patterns (e.g., mentions). In addition, we will experiment with replacing the typical embedding layer with a pre-trained BERT model. We will apply multiple convolutions using different window sizes (number of words). The CNN outputs, one per comment, can be combined by one output layer – namely by either a max layer or by a sorting layer, similar to the one proposed by us in [40]. The sorting layer is a generalization of a max layer, which takes all comments into account by probability of relation to SD.

We plan to investigate multiple variants of our proposed CNN architecture (single and multichannel), training it *both* for informative and emotional SD detection. Importantly, we will consider semantic role labeling approaches [28] to better capture language of SD, potentially decrease variance and improve predictions. To do so, we will not only investigate advanced mechanisms for **learning the language of SD**, but we will pay considerable attention to the **semantic context** that surrounds this behavior.

Context of SD - secondary features We will annotate datasets according to the following dimensions.

Topic and Context. We will include semantic contextual features that can affect the rate of SD, such as: topics of discussion, framing language, and importantly, the behavior (i.e., SD) of peers. Peer effects on privacy decision-making are widely accepted, but the nature of peer influence on SD is an open question.

- *Sentiment.* We will include annotation of "sentiment" to better contextualize SD. In recent work, we have successfully used NLTK's Vader as well as LSTM-based language models for sentiment classification [39]. Other feature-based and embedding-based methods for sentiment analysis will be considered.
- *Privacy.* We will quantify how *unique* SD utterances in a conversation lead to heightened privacy risks, and how selected sensitive topics influence the rate of SD. One promising approach is to exploit topic similarity analysis, and calculate unique utterances of SD within a conversation. For example, we can treat the output of an individual in a group as a random variable and compute information theoretic metrics (e.g., mutual information or Kullback Liebler) between that random variable and the group output (without that individual) to measure how "surprising" that individual is with respect to the group. We will refine this method for selected categories of disclosure, focusing on informational categories.

Finally, we will explore cultural factors surrounding the privacy crisis experienced during this pandemic. In particular, we will collect, analyze, and compare US-based data with data from Italian social and mainstream media, in collaboration with researchers from Italy. Italy was one of the first countries severely impacted by COVID-19. For data collection, we will select a mix of international social platforms (e.g. Twitter, Reddit Italy) as well as national outlets (Repubblica) where users' commentaries are popular. We anticipate that this effort will produce a smaller dataset with a similar timeline (February-May 2020).

IV. Early findings

Recently, our team has suggested that, in addition to risks posed by contact tracing protocols, a parallel set of privacy risks are emerging associated with information sharing through social media. Our work focuses on the phenomenon of so-called self-disclosure, or the voluntary revealing of personal information , e.g. motives, desires, feelings, thoughts, experiences, in exchange for social reward, i.e., engagement, support, legitimacy. Through analysis of 53,557,975 Coronavirus-related Tweets, we studied the shift in prevalence of self-disclosure on Twitter from January through August 2020 (see Figure 1) [4].

Between January 21 and March 11, 2020, the average daily percentage of self-disclosing Tweets was 14.63%; from March 12 through May 15, it rose to 18.89%. This change in activity coincides with an escalation in severity and increased global awareness of the crisis, with the World Health Organization (WHO) officially classifying Coronavirus as a pandemic (March 11). Current events coincident with observed changes in the rate of self-disclosure are noted on March 13 and March 19 when the United States officially declared a national stat

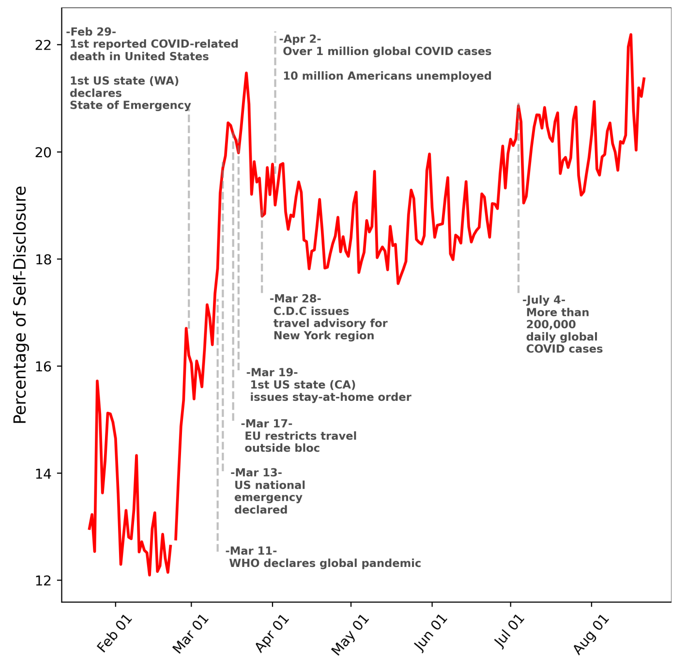

Fig. 1: Percentage of Tweets containing self-disclosure, assessed daily

emergency, and when the governor of California issued the first statewide 'stay-at-home' order, respectively. Self-disclosure activity remained high for the remainder of the dataset with an average daily percentage of 19.79% from May 16 through August 28.

When we studied the topics represented by instances of self-disclosure, we discovered an emerging focus on stay at home measures and personal, emotional experiences of the crisis. A large cluster of instances of self-disclosure in our pilot study of Twitter data in March and April are support-seeking. Words like "support" and "help" are amongst the most common in that set of Tweets. This finding is generally consistent with the thesis that social media users are using digital platforms to elicit support and stay connected during isolation.

It will be important to study whether this new approach to self-disclosure and online representation will persist beyond the pandemic. Oversharing and overexposing have been online habits of an increasingly large number of individuals online [37, 35] before the pandemic hit. Yet, the COVID crisis may have triggered a shift in the way people expose themselves online that is here to stay, in an effort to improve social connectdness and limit isolation. The implications in regards to privacy risks are yet to be explored.

V. Mitigating Risk

When we weigh the risks to privacy implicit in contact tracing protocols, we wade into complex ethical discussions about privacy-security tradeoffs. Sometimes, sacrifices in rights and liberties need to be made in service to security, in this case public health. Certainly, facing the existential threat of a global pandemic might be one such case. Although, these sacrifices should be met with scrutiny, and that is particularly the case ing crisis when fear may exert disproportionate influence.

Ultimately, our goal should be to engage measures that are maximally effective, but which we can ensure are adequately regulated and overseen. While the questions are really hard, they can at least be sketched out for concrete choices we can make – whether to engage in contact tracing and selecting amongst viable approaches to doing so.

The risks associated with self-disclosure on social media are substantively different. There are no clear set of options in front of us. There is no clear decision to make. Existing, voluntary participation in these platforms is already widespread. Government regulation of Facebook and Twitter is a topic of continual discussion, but even aside from the social media giants, there are no shortage of opportunities for individuals to share personal information online through ad hoc engagement with websites, from comments on digital news content to product reviews.

Efforts to mitigate risk will likely need to include user education and awareness. For some, learning will be experiential. Privacy risks are difficult to measure and highly contextual. Individual disclosures may seem innocuous, but increasingly sophisticated inference algorithms running over aggregations of innocuous bits of personal information can translate to deep personal insights. For many, intangible and unspecific future harm is outweighed by the immediate gains of sharing. In the extraordinary circumstances we currently share, this may be especially true.

Acknowledgments

Work from Dr. Squicciarini, Dr Rajtmajer and Umar was partly funded by the National Science Foundation under Grant Number 2027757.

References

[1] Chandan Akiti, Sarah Rajtmajer, and Anna Squicciarini. Contextual representation of self-disclosure and supportivenss in short text, 2020.

[2] Azy Barak and Orit Gluck-Ofri. Degree and reciprocity of self-disclosure in online forums. *CyberPsychology & Behavior*, 10(3):407–417, 2007.

[3] Caroline Bettinger-Lopez. A double pandemic: Domestic violence in the age of covid-19. *Council on Foreign Relations*, 2020.

[4] Taylor Blose, Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer. Privacy in crisis: A study of self-disclosure during the coronavirus pandemic. *arXiv preprint arXiv:2004.09717*, 2020.

[5] Dalvin Brown. Churches are closing doors, live streaming services for congregants avoiding coronavirus. *USA Today*, 2020.

[6] Stephen Cavey. Hipaa compliance in the age of covid-19. https://www.managedhealthcareexecutive.com/view/hipaa-compliance-in-the-age-of-covid-19.

[7] Cynthia Chew and Gunther Eysenbach. Pandemics in the age of twitter: content analysis of tweets during the 2009 h1n1 outbreak. *PloS one*, 5(11), 2010.

[8] Page Cunningham. Texts to federal government mental health hotline up roughly 1,000 percent. *The Washington Post*, 2020.

[9] Clarissa C David, Jonathan Corpus Ong, and Erika Fille T Legara. Tweeting supertyphoon haiyan: Evolving functions of twitter during and after a disaster event. *PloS one*, 11(3), 2016.

[10] Sara L.M. Davis. Contact tracing apps: Extra risks for women and marginalized groups. *Health and Human Rights Journal*, 2020.

[11] Valerian J Derlaga and John H Berg. *Self-disclosure: Theory, research, and therapy*. Springer Science & Business Media, 1987.

[12] EJ Dickson. Okzoomer is a new dating service for quarantined college kids. *Rolling Stone Magazine*, 2020.

[13] Sindhu Kiranmai Ernala, Asra F Rizvi, Michael L Birnbaum, John M Kane, and Munmun De Choudhury. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–27, 2017.

[14] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Enric Pujol, Ingmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapidor, Narseo Vallina-Rodriguez, et al. The lockdown effect: Implications of the covid-19 pandemic on internet traffic. *arXiv preprint arXiv:2008.10959*, 2020.

[15] Oguzhan Gencoglu and Mathias Gruber. Causal modeling of twitter activity during covid-19. *arXiv preprint arXiv:2005.07952*, 2020.

[16] Haddas Gold. Netflix and youtube are slowing down in europe to keep the internet from breaking. *CNN Business*, 2020.

[17] Yaru Hao, Li Dong, Furu Wei, and Ke Xu. Visualizing and understanding the effectiveness of bert, 2019.

[18] J Brian Houston, Joshua Hawthorne, Mildred F Perreault, Eun Hae Park, Marlo Goldstein Hode, Michael R Halliwell, Sarah E Turner McGowen, Rachel Davis, Shivani Vaid, Jonathan A McElderry, et al. Social media and disasters: a functional framework for social media use in disaster planning, response, and research. *Disasters*, 39(1):1–22, 2015.

[19] Amanda L Hughes, Leysia Palen, Jeannette Sutton, Sophia B Liu, and Sarah Vieweg. Site-seeing in disaster: An examination of on-line social convergence. In *Proceedings of the 5th International ISCRAM Conference*, pages 44–54. Citeseer, 2008.

[20] Andrew Hutchinson. Twitter says user numbers are up amid covid-19 lockdowns, but warns of revenue impacts. *Social Media Today*, 2020.

[21] Jerry Jackson. Doing virtual yoga during the coronavirus pandemic. *The Baltimore Sun*, 2020.

[22] Kokil Jaidka, Iknoor Singh, Jiahui Lu, Niyati Chhaya, and Lyle Ungar. A report of the CL-Aff OffMyChest Shared Task: Modeling Supportiveness and Disclosure. In *Proceedings of the AAAI-20 Workshop on Affective Content Analysis*, New York, USA, 2020. AAAI.

[23] Meg James. Fcc oks verizon request for more capacity as network use spikes. *Government and Technology*, 2020.

[24] Manon Jurgens and Ira Helsloot. The effect of social media on the dynamics of (self) resilience during disasters:

A literature review. *Journal of Contingencies and Crisis Management*, 26(1):79–88, 2018.

[25] Heather Kelly. It's 'quarantini' time. people say cheers to video happy hours when they're stuck at home. *The Washington Post*, 2020.

[26] Soo-Min Kim and Eduard Hovy. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, pages 1–8. Association for Computational Linguistics, 2006.

[27] Ashley Kirzinger, Liz Hamel, Cailey Munana, Audrey Kearney, and Mollyann Brodie. Kff health tracking poll - late april 2020: Coronavirus, social distancing, and contact tracing. *Henry J Kaiser Family Foundation*, 2020.

[28] Chenliang Li, Haoran Wang, Zhiqian Zhang, Aixin Sun, and Zongyang Ma. Topic modeling for short texts with auxiliary word embeddings. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 165–174, 2016.

[29] Sijia Li, Yilin Wang, Jia Xue, Nan Zhao, and Tingshao Zhu. The impact of covid-19 epidemic declaration on psychological consequences: a study on active weibo users. *International journal of environmental research and public health*, 17(6):2032, 2020.

[30] Siyue Li, Bo Feng, Na Li, and Xuan Tan. How social context cues in online support-seeking influence self-disclosure in support provision. *Communication Quarterly*, 63(5):586–602, 2015.

[31] Phoenix KH Mo and Neil S Coulson. Exploring the communication of social support within virtual communities: a content analysis of messages posted to an online hiv/aids support group. *Cyberpsychology & behavior*, 11(3):371–374, 2008.

[32] Paul Mozur, Raymond Zhong, and Aaron Krolik. In coronavirus fight, china gives citizens a color code, with red flags. *The New York Times*, 2020.

[33] Alex Schultz and Jay Parikh. Keeping our services stable and reliable during the covid-19 outbreak. *Facebook*, 2020.

[34] Kelly Servick. Covid-19 contact tracing apps are coming to a phone near you. how will we know whether they work? *Science*, 2020.

[35] Frances Shaw, Jean Burgess, Kate Crawford, and Axel Bruns. Sharing news, making sense, saying thanks: Patterns of talk on twitter during the queensland floods. *Australian Journal of Communication*, 40(1):23–40, 2013.

[36] Paul Shinkman. State department: China working with russia to spread coronavirus disinformation. *US News and World Report*, 2020.

[37] Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer. Detection and analysis of self-disclosure in online news commentaries. In *Proceedings of The Web Conference (WWW)*, 2019.

[38] Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer. Detection and Analysis of Self-Disclosure in Online News Commentaries. In *The World Wide Web Conference on - WWW '19*, pages 3272–3278, New York, New Yo USA, 2019. ACM Press.

[39] Jenq-Haur Wang, Ting-Wei Liu, Xiong Luo, and Long Wang. An LSTM approach to short text sentiment classification with word embeddings. In *Proceedings of the 30th Conference on Computational Linguistics and Speech Processing (ROCLING 2018)*, pages 214–223, Hsinchu, Taiwan, October 2018. The Association for Computational Linguistics and Chinese Language Processing (ACLCLP).

[40] Haoti Zhong, David J Miller, and Anna Squicciarini. Flexible inference for cyberbully incident detection. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 356–371. Springer, 2018.