Spotlight Tweets: A Lens for Exploring Attention Dynamics within Online Sensemaking During Crisis Events

2

KAITLYN ZHOU, Stanford University
TOM WILSON, KATE STARBIRD, and EMMA S. SPIRO, University of Washington

In this article, we introduce the concept of a *spotlight social media post*—a post that receives an unexpected burst of attention—and explore how such posts reveal salient aspects of online collective sensemaking and attention dynamics during a crisis event. Specifically, we examine the online conversation surrounding a false missile alert in Hawaii in January 2018. Through a mixed-methods analysis and visualizations, our research uncovers mechanisms that lead to rapid attention gains, such as *spotlighting*—when a user with existing influence confers attention by sharing others' content with their audience. We highlight how spotlight social media posts (specifically *spotlight tweets*) are distinct from other heavily shared content and that they offer insight into previously overlooked patterns in information exchange. We additionally reveal that attention dynamics may alter the social position of spotlight post authors immediately afterward (and possibly in the long term). We argue that spotlight social media posts offer a productive window for understanding online collective sensemaking, and we discuss how this can inform social media platform design and serve as a basis of future research.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; **Empirical studies in HCI**; **HCI theory, concepts and models**;

Additional Key Words and Phrases: Social computing, virality, information diffusion, social network influence

ACM Reference format:
Kaitlyn Zhou, Tom Wilson, Kate Starbird, and Emma S. Spiro. 2023. Spotlight Tweets: A Lens for Exploring Attention Dynamics within Online Sensemaking During Crisis Events. *ACM Trans. Soc. Comput.* 6, 1-2, Article 2 (June 2023), 33 pages.
https://doi.org/10.1145/3577213

1 INTRODUCTION

At 8:07 in the morning of January 13, 2018, a warning of an incoming missile was broadcast to people in Hawaii through that state's official emergency alert systems. Millions of people received mobile phone messages advising them to seek shelter and warning them that it was "not a drill."

This article is based upon work supported by the National Science Foundation under grants 1749815 and 1715078. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. K. Zhou was supported by a Stanford Graduate Fellowship.
Authors' addresses: K. Zhou, Computer Science, Stanford University, Stanford, CA; email: katezhou@stanford.edu; T. Wilson, Human Centered Design and Engineering, University of Washington; email: tomwi@uw.edu; K. Starbird, Center for an Informed Public, Human Centered Design and Engineering, University of Washington; email: kstarbi@uw.edu; E. S. Spiro, Center for an Informed Public, The Information School, University of Washington; email: espiro@uw.edu.

2469-7818/2023/06-ART2 $15.00
https://doi.org/10.1145/3577213

As is now common in crisis events [8, 25, 26, 55], people (in Hawaii and elsewhere) "converged" online to seek and share information—and to try to make sense of the alert and the impending disaster. If you were one of the numerous people who tuned into the online discourse about this event, you may have seen this tweet:

> "I just called up civil defense about the ballistic missile threat to #hawaii they said that it was a mistake." (8:12 AM)

You probably wouldn't have recognized the author—they were a relatively obscure Twitter user with just 35 followers at the time of this post. But their audience on this day turned out to be much larger. Their tweet, which appears to be the first social media post to explicitly correct what turned out to be a false missile alert, received 372 retweets. This number is orders of magnitude higher than the author's typical retweet rate (<1 retweet per tweet), showing a rapid convergence of (newfound) attention onto this tweet and, consequently, this account. We estimate that this tweet could have been seen by as many as 550,000 people within a few hours as a result of these reshares.[1] This surge in attention had long-lasting effects; the tweet's teenage author was written up in news articles and even made an appearance on a late night television show [47].

But how and why did the online crowd come to confer unexpectedly outsized attention onto this tweet? What features of the tweet, the author, and the underlying social network factored into this sudden growth? What actions—by the author and others—shaped its trajectory? How might one identify other cases of outsized attention dynamics? And what design and research insights can we learn from these interacting mechanisms?

This research looks at the intersection of attention dynamics on social media (e.g., platform affordances, network structures, user behaviors) and collective sensemaking during crisis events. Collective sensemaking is a process through which people attempt to understand, assign meaning to, and make decisions about disruptive events [61, 82]. This process once took place as physical gathering at the site of a crisis [18], but is now occurring online as people—including those locally affected as well interested spectators around the world—converge onto the virtual scene [27, 35, 49, 56]. And this online convergence is intersecting with the information dynamics of online platforms. Previous research has investigated why certain authors receive additional attention during crisis events—for example, because they are an eyewitness or because they have local or other relevant expertise [38, 71, 84]. Our work expands this view to look at how these authors—and their specific tweets—accumulate that attention. How does a tweet receive 372 retweets when the author would typically get one or none? How can we unpack those 372 retweets to understand mechanisms that confer attention? And if we can do this, what are the design implications for social platforms and directions for future research?

To answer these questions, in this article we introduce the concept of a *spotlight social media post*, a post like the preceding example that receives an unexpected burst of attention—operationalized here as an order of magnitude more than is typical for that user. In the context of Twitter, we refer to these posts as *spotlight tweets*. Focusing on Twitter activity during the 2018 Hawaii False Missile Alert, we identify a set of 212 spotlight tweets and explore what they reveal about the underlying dynamics that shape how attention is distributed during a mass convergence event. Our case study allows us to build a comprehensive understanding of spotlight tweets, and in future work it is our hope to study how these spotlights emerge and behave across various other crisis events.

We employ a mixed-method approach, using visualizations to guide in-depth qualitative analysis to understand how specific tweets gained attention over time and fit into the collective effort to

[1] Estimates of audience are difficult. We follow the method from Arif et al. [1] to estimate audience size using the follower count of users who retweet the content.

make sense of the missile alert, and integrating quantitative analyses to better understand what features of tweets and authors were most associated with spotlight tweets overall.

Our research makes several contributions. First, we demonstrate that spotlight tweets provide a productive window into attention dynamics—revealing what information the crowd finds particularly valuable during the event as well as how the crowd comes to recognize that value. We create an annotated cumulative retweet and quote graph[2] with audience indicators—as a method for investigating how tweets garner attention over time, specifically highlighting the intermediary role of large-follower accounts. We identify distinctions between spotlight tweets and other highly retweeted tweets (which we call *popular tweets*), surfacing insights that may be missed when researchers study the highest retweeted tweets based on the retweet distribution of the crisis event. Among other findings, we note that spotlight tweets were more likely (1) to be authored by users who were located in Hawaii at the time of the event, (2) to spread rather than correct the missile alert, (3) to include an attached screenshot of the alert, and (4) to contain a personal account (either first-hand or second-hand) of someone who had experienced the alert and its immediate aftermath. We also demonstrate that many spotlight tweets were retweeted, prior to their surge in growth, by one or more users with a large audience. We describe how *spotlight operators* (influential accounts that selectively amplify other's content) perform the role of distributing attention, likely shaping the online sensemaking process that occurs during crisis events. To demonstrate the applicability of the concept to other conversations, we explain how spotlight tweets help understand information dynamics in two other cases studies. Finally, we conclude with several design and research insights gained from studying spotlights tweets. It is our hope that these insights will help designers build more socially translucent social media platforms and will equip future researchers with a better understanding of the impacts of attention dynamics during crisis events.

2 RELATED WORK

2.1 Sensemaking During Crisis Events

The use of social media during crisis events is a well-established phenomenon with platforms increasingly used in crisis communication by local responders and government agencies [7]. Social media also provides a venue where people, including locally affected individuals and a global audience, converge to view, contribute, and share event-related messages [21, 54]. Crises, such as natural disasters or terrorist attacks, are characterized by high levels of uncertainty [11]. Historically, this uncertainty was often attributed to a lack of information, especially from official or authoritative sources [13, 16], but with the advent of social media there is also the contradictory problem of information overload—an abundance of information [24, 48, 62], some of which may be conflicting or inconsistent. One role of social media during crises is its use in collective sensemaking, a process through which people and communities sift through this information and attempt to resolve some of the uncertainty—to understand what is happening, why, and what they should do about it [22, 49].

Sensemaking during disaster events precedes social media. For example, Fritz and Mathewson [18] spoke of "informational convergence"—that is, how people converged onto disaster-affected areas, both physically and via available communication channels (e.g., telephones and mail), to figure out what was going on. During a crisis event, people can experience a sudden loss of meaning brought on by the disruption [82]. In response, a collective process can take shape—driven by the need to reestablish order and meaning in the moment. Weick [82] termed this process

[2]Our annotated retweet graph contains both retweets and quote tweets (formerly known as retweets with comment).

sensemaking, which he described as a collective process to reestablish order and make retrospective sense of what has occurred. This kind of sensemaking matches the description by sociologist Shibutani [61] of a highly distributed problem-solving activity that functions to fill knowledge gaps and make sense of an emerging situation. In the context of a crisis situation, Shibutani describes the formation of a public, an unorganized group of individuals lacking conventionally defined roles that are brought together in a dynamic situation through a common interest.

The advent of computer-mediated communication, and later social media, has seen a surge in research about collective sensemaking online and particularly its relationship to the spread of rumors and misinformation [50, 69]. Studies have highlighted the ways in which Twitter has been appropriated for communicating during times of crisis [38, 40]. However, patterns of collective information processing on Twitter are similar to that observed offline in prior crisis events, and a lack of clear source, personal involvement, and anxiety are significant drivers of sensemaking activities [48]. This growing body of work suggests that during crisis events, the online collective sensemaking process is shaped by users posting, sharing, adding information, and deriving meaning from event-related social media posts. Understanding why and how some of this information may stand out against other information—in other words, the information that receives more attention—is key to sensemaking because information that is visible, accessible, and salient will play a larger role in shaping meaning.

2.2 Attention Dynamics and Social Media Platforms

Simon [65] highlighted the scarcity of attention, positing that the wealth of information available was being met with a dearth in the attention that the information consumes. Sterling [75] referred to this as an information economy in which "everything is plentiful—except attention." Attention is not equally distributed. Instead, it is an accumulated advantage phenomena: "the rich get richer and the poor get poorer." This idea, also known as the Matthew effect, can play a role in how attention is distributed through a network. The Matthew effect was traditionally used to describe the distribution of credit in academic networks, with credit often assigned to eminent scientists over unknown researchers despite the true distribution of labor behind the project [44]. In the context of social influence, experiments have shown that individual behaviors toward music and books follow the preferences of those before them [60, 66]. The principle of accumulated advantage has also been observed in social networks, where new nodes are more likely to connect to others that are already well connected in a process known as preferential attachment [32].

Attention is a high value resource in the ad-based internet economy [86]. The ability to capture attention is particularly coveted by businesses, politicians, and entertainment agencies because it can be channeled to someone or something else, and ultimately monetized [20]. Attention can also be used to gain influence—a user's ability to propagate information through a network [59]. The more one's audience is able to propagate their information, the more attention they capture, and the more likely their content will trend in social networks and become a part of a public agenda. Some online opportunists have recognized that crises and other breaking news events can be opportunities to gain attention [43].

In the context of social media, one behavior that may follow the principle of accumulated advantage is retransmission of content. On Twitter, this is most often done through the retweet affordance. Previous research has explored factors that underlie and motivate retweeting behaviors. Boyd et al. [4] argue that one of the reasons for retweeting is to "to amplify or spread tweets to new audiences." When deciding what to retweet, users choose to retweet information that would be of interest to their imagined audience [4]. Similarly, Welch et al. [11] describe reshares as endorsements of content or its author, although many Twitter users explicitly resist this perception, claiming "RTs =/= endorsement" in their profiles. Other researchers [2, 31, 76, 78, 89] have tried to predict which tweet features correlate with retweet counts, attempting to identify what type of content will be of interest to a larger audience. Whether endorsement, appropriation, or simply amplification, retweets function to shape how information and attention flow on the platform, as one user can choose to direct the attention of their audience to another user's content.

Our research looks at retweets (including retweets with comment, i.e., quote tweets) as both a shaping force and a reflective measure of attention on Twitter. In this article, we examine how retweets function to focus, confer, and distribute attention—and how those attention dynamics intersect with the collective sensemaking process that occurs during a crisis event.

2.3 Spotlight Social Media Posts

In particular, we focus on cases when a user's tweet receives an abnormally high amount of attention (based on their typical retweet rate) in the form of retweets. We introduce the term *spotlight tweet* to characterize these incidents—where a user gains notable attention to their content and platform profile in an unprecedented way, often coinciding with a few retweets from highly influential users. The focus on spotlight tweets rather than the most retweeted tweets during a crisis event is intentional, as we are not studying which tweets received the most attention but rather which tweets gained attention in proportion to the user's previous average. We then compare spotlight tweets with the most retweeted tweets, often from users with large followings such as politicians, journalists, and news organizations. These attention-gaining events likely have long-term consequences on the underlying structure of the network, as the accounts that authored spotlight tweets can themselves grow in visibility (through increases in follower count). Previous research has documented this effect for accounts of local responders during crisis events [67]. Our work explores these effects for any account that receives outsized attention during a particular crisis event. Expanding upon previous work on information cascades [14, 28, 58] and identifying content and user metadata that may contribute to retweet rates during a crisis [46, 78], our work takes an in-depth look at how these tweets garner that attention and explores how those attention changes could shape the collective sensemaking process online.

3 HAWAII FALSE MISSILE ALERT EVENT DESCRIPTION

This study primarily focuses on a specific crisis event—the Hawaii False Missile Alert, which occurred on January 13, 2018, when a missile alert was sent to the residents of and visitors to Hawaii via the Emergency Alert System and Commercial Mobile Alert System. The alert went out through television, radio, and mobile phones. Although recipients would later learn that the alert was sent in error, for nearly 40 minutes many people on the islands of Hawaii thought that a missile attack was imminent. For hours afterward, the incident held the attention of a global audience.

For researchers, this case offers a unique window into how affected populations respond to imminent hazard alerts—and how global audiences interact in an increasingly politicized online discourse around such events both during and after the event itself. This perspective is supported by a history of work in the social sciences demonstrating that crisis events produce complex and uncertain information environments that are ripe for misinformation, political propaganda, and disinformation spread [25, 61, 69, 91].

The Hawaii False Missile Alert sent at 8:07 AM Hawaii DST (18:07:00 UTC) read:

> "BALLISTIC MISSILE THREAT INBOUND TO HAWAII. SEEK IMMEDIATE SHELTER. THIS IS NOT A DRILL."

At 8:12 AM, the alert was canceled, meaning alerts would not be sent to phones that had yet to receive the message (either due to network connections or devices being turned off) [37]. However, the public was not initially given any information explaining the alert or why it was canceled. Then, 38 minutes later, at 8:45 AM, an official correction was sent out:

> "There is no missile threat or danger to the State of Hawaii. Repeat. False Alarm."

We focus on the first 2 hours of Twitter activity following the initial false alert (8:07 AM to 10:07 AM Hawaii DST). This data window covers several key features of the event: the initial alert, early unofficial corrections, sensemaking about the nature of the alert (including whether it is accurate or not), official corrections, and the beginnings of politicized conversations about who is at fault.

4 DATA DESCRIPTION

Event Tweet Data. Our Hawaii False Missile Alert event data was obtained from Twitter's Historical PowerTrack via DiscoverText's Sifter product.[3] Historical PowerTrack provides access to public tweets from an archive of the "firehose"—the stream of publicly available Twitter data at a specific point in the past [36].

Our request captured all public tweets (i.e., tweets not deleted by the user or authored by an account since subject to suspension or set to protected (private) at the time of request) posted between 08:07:56 AM and 10:07:00 AM Hawaii DST (18:07:56 to 20:07:00 UTC) on January 13, 2018, that contained either one or more of the keywords ("hawaii," "missile," "alert," "false alarm," "falsealarm") in the tweet body, and/or references to locations in the state of Hawaii[4] in the

- tweet body,
- user's location, or
- user's biography.

This request resulted in a total of 285,557 public tweets (including 97,051 original[5] tweets) with 150,442 distinct users participating. We archived this dataset on February 1, 2018. Any tweet deleted or otherwise removed from public view before this date would not appear in the data. Since the event took place in Hawaii, we can roughly differentiate between local populations and global audiences using the UTC-offset information in tweet metadata.

A prior comparison of Historical PowerTrack to Twitter's Search and Streaming API on four distinct events found PowerTrack data to offer the most complete set of data, because it is not subject to constraints such as rate-limiting [80], which can be encountered during collections on events with high tweet volumes, and can include emergent terms such as *false alarm* that may have been hard to predict at the time of collection.

User Account Historical Tweet Data. In addition to the event tweet data, we also capture data about each unique user account's history of posts (i.e., their timeline of tweets). This historical data is needed to estimate a prior-to-event average retweet count for their content. We use this baseline engagement data to identify spotlight tweets during the event. We used the Twitter API on November 21, 2018, to query each user account's history of posts, up to a maximum of 3,200 of the most recent posts. Since this collection was made months after the event, it is possible that some tweets were deleted during this time and these deleted tweets would not be captured. We also note that for some extremely active users, we are unable to collect pre-event tweets using this method; in other words, the historical 3,200 tweets do not extend prior to the event observation period. In some cases, these user accounts are removed from our analysis. In others, we were able

[3]The DiscoverText website is available here: https://discovertext.com/.
[4]List of locations in the state of Hawaii as listed on the Wikipedia article, "List of places in Hawaii."
[5]Original tweets include quote and reply tweets and exclude retweets.

to estimate pre-event retweet rates using complementary user data shared from Pushshift, which we discuss in more detail in the following.

Limitations. In both our datasets, we are limited to tweets that have not been deleted at the time of collection. Previous research around deleted tweets has shown that a potentially large percentage (over 50%) of deleted tweets contain negative sentiment, content on relationships, and cursing [90]. In the case of the Hawaii event tweet dataset, we recognize that our research maybe be limited as deleted tweets will be missing. This challenge may also affect the user account historical data, but since we use this data to estimate the user's prior-to-event average retweet count, it is less likely to be significantly affected by missing data.

5 METHODS OF ANALYSIS

5.1 Defining and Operationalizing Spotlight Tweets

Seeking to better understand the phenomenon of attention gain, we introduced the concept of a *spotlight tweet*—a tweet that rises in visibility in an unexpected way. We propose that spotlight tweets can be identified through metrics that account for the user's baseline audience engagement and detects a significant upward shift in the attention they receive for a specific post. In the Hawaii False Missile Alert dataset, we operationalize spotlight tweets as tweets that have retweet counts that are 10 times greater than the user's previous average.[6,7] Given that many users have a below 1 retweet average, we want to sure that there is a minimum threshold of attention that a user has gained to be considered a spotlight tweet. Here, we set the minimum number of retweets to be 10.

We scope our analysis to a set of users who (1) actively tweeted at least twice in the first 2 hours and (2) had one of their original tweets retweeted or quoted[8] at least twice. Scoping the study population results in a tractable set of users (both in terms of additional data such as user histories that needs to be queried from Twitter and of a set of users that can be qualitatively explored). This scoping left a set of 3,624 unique users. For all users whose timeline query (as described in Section 4) retrieved at least 50 tweets prior to the Hawaii False Missile Alert event (January 13, 2018), the mean number of retweets received per message per user was computed (i.e., the prior-to-event average retweet count per user). Only tweets from users whose prior-to-event average retweet count could be calculated with relatively low variance are considered when identifying spotlight tweets. This baseline information is necessary to ensure that spotlight tweets are departures from normal attention received. This results in a set of 1,203 unique users. Baseline engagement averages are used to categorize tweets in our Hawaii Missile Crisis dataset from these authors as one of the following three categories:

- *Spotlight* (S) tweets are event-related tweets with more than 10 retweets and a retweet count that is at least 10 times greater than the user's baseline average retweet count (n = 212).
- *Popular* (P) tweets are event-related tweets that are among the top-1,000 most retweeted tweets in the dataset (this resulted in a set of 980 tweets each with at least 34 retweets). We limited the set of popular tweets to only tweets from users for whom we could calculate pre-event averages for (using either the Twitter REST API or data provided by

[6]We find that the distributions of retweet counts of users are not normally distributed, making the use of standard deviations inappropriate.

[7]The retweet count of tweets continues to change, as many are still public on Twitter. For the purpose of our study, the retweet counts used were queried from Twitter between May 15 and 16, 2019.

[8]A quoted tweet, formerly known as a "retweet with comment," is an original tweet where a user embeds another tweet in their tweet and adds addition content.

Pushshift). This allows us to ensure that these popular tweets (n = 413) are not spotlight tweets.[9]

- *Other* tweets are all other event-related tweets in our dataset that do not meet the definition of one of the previous categories (n = 95,972).[10]

Our specific method of calculating a user's prior-to-event average retweet count might introduce bias against tweet authors who are very active (due to rate limits from the Twitter Search API). For highly active users, we are unable to obtain observations of behavior prior to the event case studied here and therefore unable to compute a baseline engagement rate. In an effort to better understand this bias, we used two additional methods to identify baseline attention received and thus spotlight tweets during the event studied here:

(1) We compute baseline attention with the user's event and post-event average retweet count as a proxy for their prior-to-event average.
(2) We used an third party tool, Pushshift, that makes use of API architecture to collect the 500 most recent tweets leading up to the event for our set of active and early engaged users.

Using these two alternative methods, the number of spotlight tweets identified changed slightly; however, the rates at which these additional spotlight tweets used Twitter affordances (quotes, replies, hashtags, location names) were not significantly different from the data used in the study. This suggests that despite the possibility for bias, in practice, we would expect our findings to remain consistent given a more complete dataset.

5.2 Mixed-Method Approach to Analysis

To understand who spotlight tweeters are and how/why their event-related tweets garnered attention, we employed a mixed-method approach, building upon previous methodological innovation in crisis informatics [53] and the study of online misinformation in the crisis context [42]. This approach blends descriptive quantitative and visual analysis to get a high-level view of patterns and anomalies in the data with close-up qualitative investigation of the data that generate those patterns and anomalies. The methodological process is iterative, in that we move back-and-forth across the different levels of investigation. It is also a grounded, interpretative process—following Charmaz and Belgrave [10]—where insights emerge from this multi-level engagement with the data. This methodology aligns well with the "computational grounded theory" approach described by Nelson [45]. All together, this process integrated high-level pattern analysis and visual interpretation, deep qualitative content analysis including the manual inspection of hundreds of tweets and their patterns of retweets (including analyzing both who retweeted and when), and quantitative analysis to evaluate emergent hypotheses and quantify specific trends in the data.

In this study, we applied these techniques to the investigation of spotlight tweets—looking at how they gained attention through cascades of retweets. Retweets often occur through tree structures, where downstream exposures are mediated by intermediate accounts. In other words, I may retweet something from an account I do not follow (Account A) because I saw it through a retweet by an account I do follow (Account B). However, in practice, the digital traces of behavior one can observe from Twitter do not reveal these tree structures; instead, my retweet is assigned in the tweet metadata (and appears in users' views) as a retweet of Account A—with no mention of the mediating role of Account B. For this reason, analyzing retweet counts alone may obscure

[9]This set is smaller than the original 980 through filtering out authors for whom we do not have pre-event averages and through ensuring they are not spotlight tweets.
[10]Note that the total number of tweets does not quite add up to 97,051 tweets since popular tweets that could not be verified as non-spotlight tweets are not included.

some of the more interesting dynamics regarding how a piece of information gains attention and spreads—that is, what other accounts may have played a role.

For example, previous work has demonstrated that online crowds tend to amplify the content of local individuals (i.e., those physically located in the area of impact) more than other content [38, 71, 73], but it is unclear if each of those retweeting accounts identified those locals on their own, or if they relied to some extent on the "work" of others to call their attention to accounts with "local authority."

5.3 Visually Guided Qualitative Analysis of Spotlight Tweets

To better understand these dynamics, we use a visualization technique referred to as an annotated cumulative retweet and quote graph that overlays user information, in this case follower counts, onto a temporal graph. This graph, featured throughout the article, builds upon previous efforts (see the work of Arif et al. [1]) that attempt to shed light on the role that large audience accounts play in shaping the trajectories of information cascades. Here, we used the graphs as interpretive artifacts, to guide qualitative analysis of a sample of different types of tweets (spotlight tweets, popular tweets, other tweets) to understand how attention dynamics shaped the spread of specific tweets, and how these tweets shaped the broader sensemaking process around the false missile alert. In particular, these graphs allowed us to identify and explore the role of intermediate accounts—accounts that we term *spotlight operators*—that focused our attention toward what became spotlight tweets.

5.4 Quantitative Comparison of Metadata Features of Spotlight, Popular, and Other Tweets

Informed by insights that emerged through our visual-qualitative analysis, we then conducted quantitative analyses on tweet metadata to identify specific features of tweets or accounts that might have made them more likely to become spotlight tweets in general across our whole datasets. We calculated correlations between tweet types (spotlight, popular, all other) and author characteristics like location (inferred from time zone information), tweet formats (original, retweet, quote tweet), and tweet features (links, hashtags, location names). In the following, these analyses are described in more detail in their relevant sections.

5.5 Systematic Qualitative Coding and Comparison of Content Features of Spotlight, Popular, and Other Tweets

Finally, to better understand the discursive features of tweets about the Hawaii missile crisis generally—and spotlight tweets specifically—we developed a coding scheme (grounded in the data from this specific event) and applied it to random samples of spotlight, popular, and other tweets. This coding scheme emerged through a combination of open coding, where researchers identified potential codes and grouped them into broader categories, and closed coding, where researchers attempted to apply the current coding scheme, and adjusted and refined it to accommodate new insights. Throughout this process, we focused on identifying the different types of sensemaking (about the alert and its correction) that took place through tweets, as well as on surfacing specific features of tweets—such as signals of proximity to the event—that may have helped some tweets garner out-sized attention. Table 1 provides a concise overview of the coding scheme, which we explain in more detail in Section 6.

After the coding scheme stabilized, the first and last authors coded a sample of tweets that included all 212 spotlight tweets as well as random samples of 200 popular tweets and 200 other tweets. For one-third of the tweets (across samples), we used a consensus coding process, with both researchers coding each tweet and resolving disagreements through discussion. We report on the

Table 1. Overview of the Qualitative Coding Scheme

Coding Dimension	Explanation	Inter-Rater Reliability (κ)
Stance	Stance toward whether a missile is incoming. Options: correcting, spreading, neutral	.73
Sensemaking type	What kind of sensemaking information behavior does this tweet engage in? Options: informing; threat; protective action; causes, impacts, and solutions; responses; emotional sensemaking	.68
Proximity	What proximity/attachment does the tweeter have to the event? Options: primary, secondary, none	.65
Personal account	Sharing a personal account (of what they or someone else went through) Options: Yes/No	.61
Politicizing the event	Is there a political element here? Are they connecting this to a political discourse? Options: Yes/No	.68
Targeting	Does the tweet invoke some target for blame or praise? What element of the event/situation is being targeted? Who is the target?	.71/.71/.67

Cohen's kappa score was calculated between the two annotators as the inter-rater reliability value.

inter-rater reliability (Cohen's κ) for original, independent codes for those tweets in Table 1. For the remaining 400 tweets, each was coded by one of two researchers, who coded equal numbers of spotlight, popular, and other tweets to mitigate any researcher bias.

6 RESULTS

6.1 A Qualitative Look at Exemplar Tweets

To begin, we present the analysis of four exemplar tweets. These cases help explain how the alert and correction unfolded on Twitter. They also reveal characteristics of common cumulative retweet and quote signatures, including potentially meaningful differences between spotlight and popular tweets. These exemplars were selected through an iterative, exploratory process to showcase specific elements of the empirical observations of the research team. Consistent with recommendations and practices from previous literature [15, 85], we have anonymized all user names and profile pictures *except* for accounts that are of public interest as verified by Twitter[11] or for accounts that have given us explicit consent to use their tweets. They are illustrative and aid in grounding more quantitative analyses.

6.1.1 Tweet A: A Local Tweets the Alert. In the early moments of the crisis, immediately following the alert, there were many tweets relaying the fact that the author had received the alert. Tweet A, a spotlight tweet, is an example of this behavior (Figure 1). It was sent at 8:09:33 AM (Hawaii DST), about 2 minutes after the alert went out.

Tweet A was posted by Linda Nagata, a verified Twitter user whose bio signaled that they are an award winning science fiction author. Linda's account profile location was set to Maui. Significant for later analysis based on user location, Linda's UTC offset was set to −36,000, which is consistent with being in the Hawaiian-Aleutian time zone. Linda's tweet shares that they received an emergency alert on their phone indicating a ballistic missile threat to Hawaii. Linda uses a hashtag to mark Hawaii in the text.

At the time of this tweet, Linda had posted a total of 23,370 tweets according to account metadata. Across Linda's most recent pre-event 846 tweets (also collected in our study), the average number

[11]Twitter's page on verified accounts: https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts.

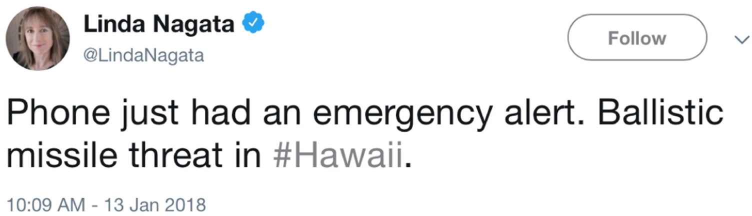

Fig. 1. Screenshot of Tweet A.

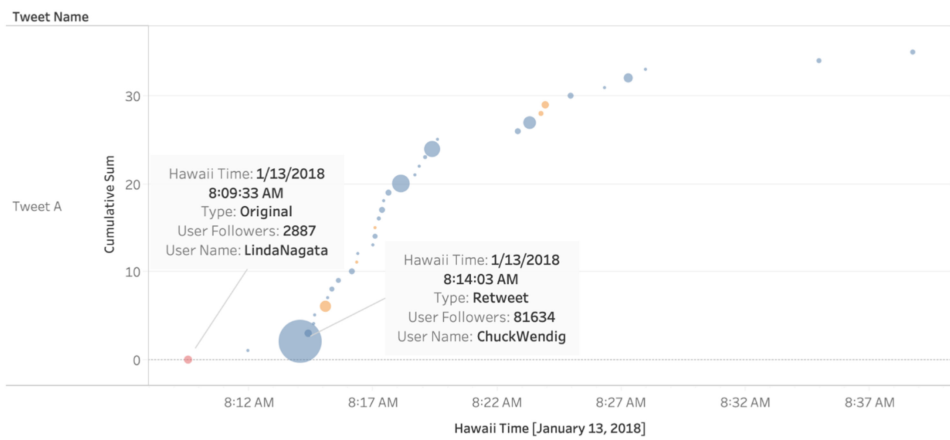

Fig. 2. Tweet A. Cumulative retweet and quote signature of Linda's tweet, marking the original tweet, and when Chuck Wendig retweeted their content. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author.

of retweets per tweet was less than 1. Tweet A, however, received 30 retweets during the 2-hour window following the initial alert. This tweet meets our criteria for a spotlight tweet.

Figure 2 shows the volume of retweets (blue dots) and quotes (orange dots) as a running sum (along the y-axis) over time (along the x-axis). Plotting the temporal signature as a cumulative sum rather than tweets-per-minute allows us to visually see inflection points of note in the retweet dynamics and to overlay additional features of the individual tweets or users to add context for these dynamics. In this case, each of the circles represents a user who retweeted Linda; the size of the circle represents the number of followers (i.e., audience) the retweeting user had at the time.

Figure 2 has several interesting features. One of those is a sharp slope change at about 8:14:03 AM. In the first 5 minutes, Linda's tweet receives only one retweet (at 8:11:56 AM). At 8:14:03 AM, the tweet is retweeted by a high-follower account, @ChuckWendig, marking the start of a sharp retweet rate increase—20 retweets in about 5 minutes. Chuck Wendig is a verified Twitter user with 81,634 followers. Chuck's profile indicates that they are a NYT-bestselling author and their location is set to "Pennsyltucky, Pennsylvania," a humorous claim that suggests that they are not local to Hawaii. Linda, who is local to Hawaii and likely directly affected by the event, initially received little attention on their post, similar to prior posts. However, this retweet from a large-follower account (who Linda was perhaps connected to due to their common profession—as authors) likely helped bring attention to the tweet, changing its trajectory and catalyzing Tweet A as a spotlight tweet for Linda.

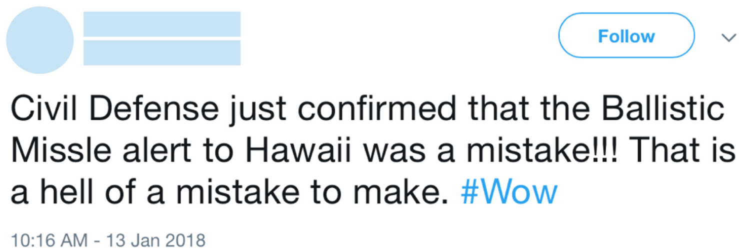

Fig. 3. Screenshot of Tweet B.

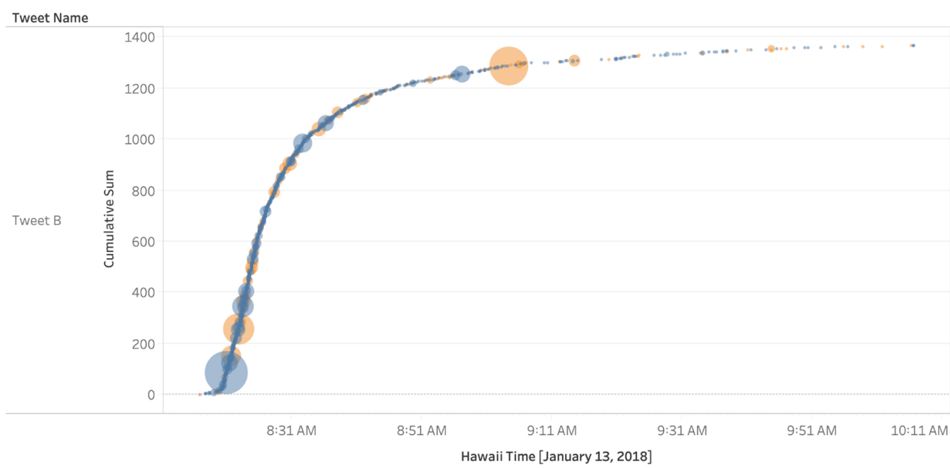

Fig. 4. Cumulative retweet and quote signature of Tweet B, where an S-shape is formed. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author.

6.1.2 Tweet B: A Local Posts an Early Correction. Users on Twitter were quick to learn and share that the missile alert was a false alert. Tweet B is one of these corrections, and it is a highly retweeted spotlight tweet (Figure 3).

The tweet was sent at 8:16:49 AM by a user in the U.S. Army, who is based in Schofield Barracks, Hawaii, according to profile content. At the time, they had 160 followers and 336 previous posts. Across their last 230 tweets, the user had an average of 0.27 retweets per tweet.

The user shares that they received confirmation from the civil defense that the ballistic missile alert to Hawaii was a mistake. The timing of this tweet is critical because although the missile alert was canceled at 8:12 AM, official messages confirming that the alert was a mistake were not sent out until late—8:20 AM on Twitter and 8:38 AM through the emergency alert system.

In the first 4 minutes, the tweet receives a total of 17 retweets (Figure 4). At 8:18:59 AM, a user with 10,289 followers retweets the tweet. At 8:19:41, a user with 8,024 followers also retweets the tweet. However, neither of these retweets qualitatively change the trajectory of tweet signature. At 8:19:44 AM, @mike_hogan who is a Vanity Fair digital producer with 11,577 followers, quotes the tweet (orange dot), adding the comment, "JFC" (Jesus Fucking Christ) to the original message (Figure 5). Over the next minute, the tweet receives several more retweets.

All of the sudden, at 8:20:18 AM, we see a dramatic shift in retweet rate. But interestingly, the first few tweets in that cascade are not from large-follower accounts. This suggests to us that the

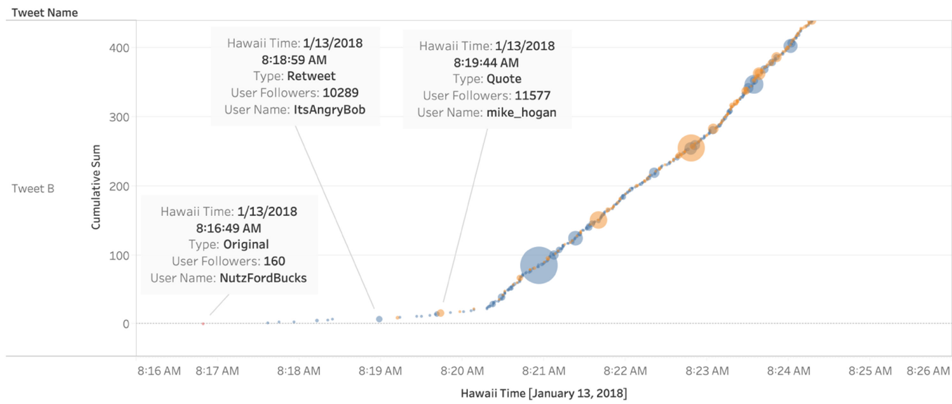

Fig. 5. Tweet B. A zoomed-in view of the initial 10 minutes is provided to highlight the key actors in play and the impact (or lack thereof) of their retweets. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author.

followers at the start of the cascade are not necessarily the only contributors to this change of trajectory for the tweet, but rather that an exogenous factor could be in play.

At 8:19:30 AM, a tweet (details are presented in the following) sent out by an official account corroborated the information that there was no incoming missile, confirming the information from Tweet B, and our data suggests it could have helped change the trajectory of this tweet—users may have been more likely to retweet it once they had seen or knew of corroboration of the information, especially since this message occurred prior to the official correction.

The tweet ultimately received 1,071 retweets. For comparison, the official Twitter account of the network ABC, which has 13M followers, shared a tweet at 8:29 AM saying there is no missile threat, and it received 1,524 retweets. The retweets to followers ratios for Tweet B and for ABC's tweet are 6.69 and 0.0001, respectively.

6.1.3 Tweet C: Quoting a Correction from a High-Profile Account. Tweet C, another highly retweeted spotlight tweet, is substantially different from the previous two examples, although its cumulative retweet and quote signature is very similar (Figure 6). Unlike Tweets A and B, this tweet was not authored by an account whose profile description indicates connection to Hawaii (rather it is Oakland, California), and it is not an "original" tweet but quote tweet (also known as a "retweet with comment") of another user—in this case, U.S. Representative from Hawaii Tulsi Gabbard. Gabbard's tweet was the first by a high-profile account and "official" representative of the state to definitely state that the alert was a false alarm. It was also a spotlight and heavily tweeted and quoted in our dataset. In subsequent sections, we look specifically at distinctions between quoted tweets and other tweets.

Tweet C was sent out by a "former newspaper guy" who at the time of the event had a relatively small audience (2,068 followers) and a relatively low activity rate (1,247 statuses). Across their last 354 tweets, the user had an average of 0.17 retweets per tweet.

The user quotes Tulsi Gabbard, but adds additional commentary to the tweet, describing the author's personal experience and the emotional reaction they had to both the alert and to learning that it had been a mistake. Though their profile description does not indicate Hawaii, their tweet suggests that they are located there, received the alert, and have been profoundly affected.

Miguel Helf (@mhelft), a journalist, is the first user to retweet this content (Figure 7). Miguel has a large-follower account with 19,979. At 8:31 AM, Matt Pearce (@mattdpearce), a verified Twitter

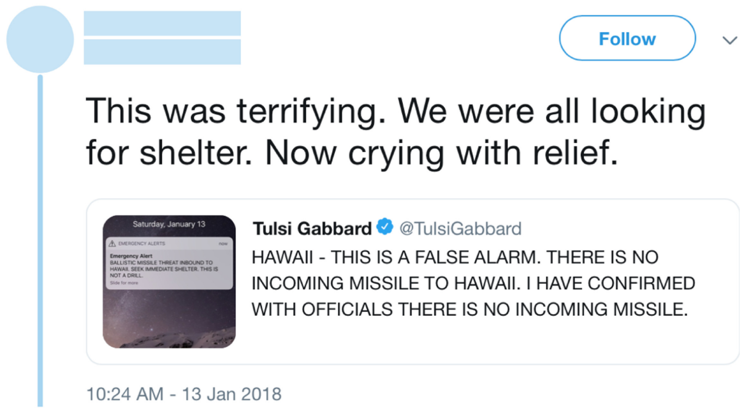

Fig. 6. Screenshot of Tweet C.

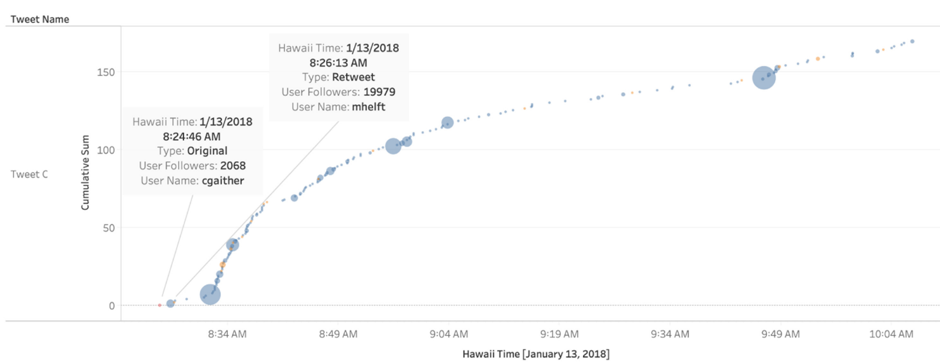

Fig. 7. Tweet C. The cumulative retweet and quote signature shows that this spotlight tweet again follows the S-shape. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author. Highlighted are two journalists with large-follower accounts.

user with 128,447 followers and a national correspondent for the *Los Angeles Times*, retweets it as well. At this point, we start to see a shift in the cumulative retweet and quote count signature with a dramatic increase in the retweet rate. Matt was the sixth user to retweet this tweet, which ultimately received 153 retweets. We suspect that although Miguel's retweet did not immediately bring significant attention, it is possible that Miguel's retweet traveled through a network of journalists, eventually catching the eye of Matt Pearce, whose retweet coincides with a change in the trajectory of the tweet.

6.1.4 Tweet D: Request to Share from a U.S. Congressmember. Similar to Tweet C, Tweet D is also a quoted tweet (Figure 8). It is classified as a popular tweet (and not a spotlight tweet). This example highlights how the cumulative retweet and quote signature can differ between a popular tweet and a spotlight tweet even if the total retweet counts are similar.

The tweet was sent out at 8:25 AM by U.S. Senator Brian Schatz from Hawaii, who has 80,815 followers. Across Senator Schatz's previous 350 tweets, @brianschatz had an average of 1,135 retweets per tweet. Their UTC offset is in the Hawaiian-Aleutian time zone, and their location coordinates are to the city of Hilo.

Senator Schatz quotes the Hawaii EMA official Twitter account, asking followers to "please retwee" this information. We know from other studies [79] that asking users to share or using the

Fig. 8. Screenshot of Tweet D.

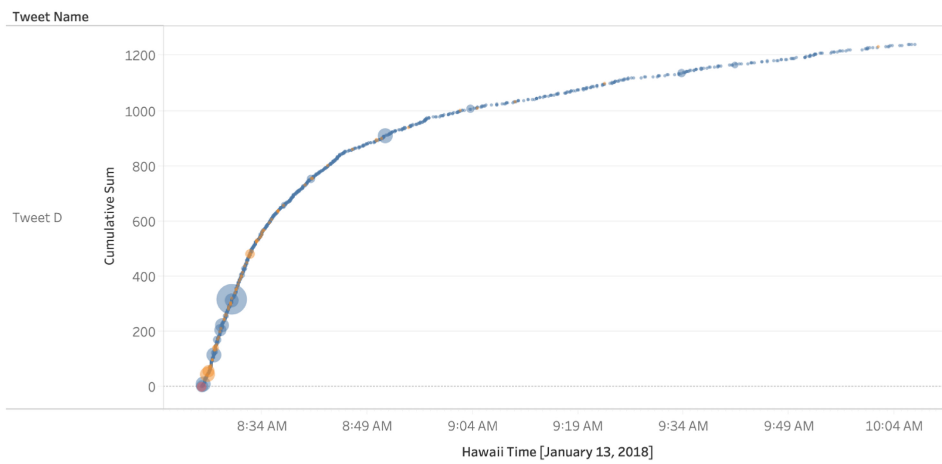

Fig. 9. Tweet D. The cumulative retweet and quote signature of Senator Schatz's tweet follows a log-like shape rather than an S-shape. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author.

words "retweet" and "please" helps increase the retweet count of a tweet. We believe that, given Schatz's position, they were trying to strategically use their position to help spread the information that the alert was a false alarm.

Different from the previous cumulative retweet and quote tweet signatures, this tweet's rate of retweet is high right from the start and continues to be retweeted at a rapid rate, eventually tapering off around 8:45 AM (Figure 9). The 2nd, 3rd, 4th, and 10th users to retweet Brian have 24,393, 150,898, 15,587, and 230,291 followers, respectively. Three of these four users represent themselves as journalists or "social media influencers" in their bios. Different from the spotlight tweets, we see that many individuals retweeting Schatz's tweets are online "influencers" who immediately pick up and retweet Schatz's message (Figure 10). In the first 3 minutes alone, there were 182 retweets. This tweet ultimately received 1,475 retweets.

Taken together, these examples demonstrate how downstream high-volume follower count accounts shape information spread. As we observe the growth of a tweet rather than simply analyzing this content and final retweet count, we reveal patterns of information exchange that otherwise would be not be found. Specifically, we demonstrate precisely how the act of spotlighting content can shape a tweet's transmission trajectory. With this in mind, a natural question to ask is: what are the unique features of tweets that get spotlighted?

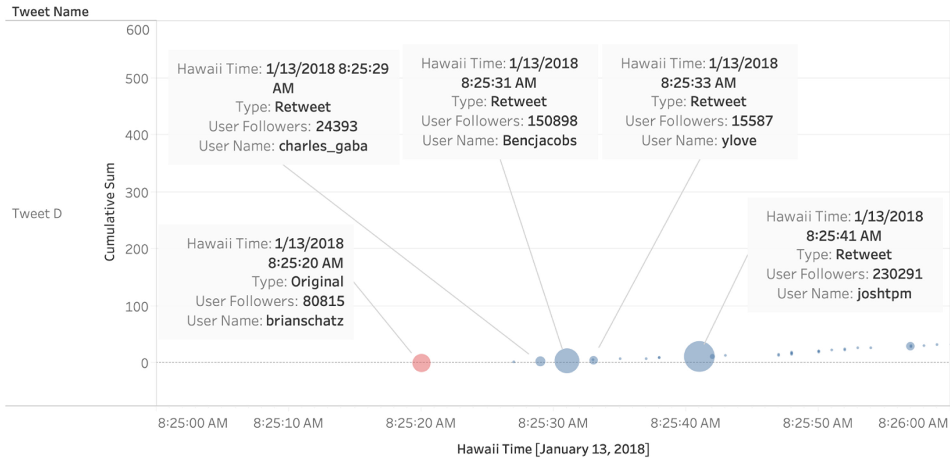

Fig. 10. Tweet D. A zoomed-in view of the initial minute is provided to highlight the early large-follower accounts immediately retweeting this information. Blue points mark retweets. Orange points represent quotes. Size of points represents the number of followers of the posting author.

6.2 Tweet Features and Tweet Types: Metadata Analysis

Attention in social media platforms can occur in many forms; attention, however, can often be very difficult, if not impossible, to measure directly. Twitter, for example, does not provide information about which tweets are read by a user. What we can do is consider a proxy for attention, and that is engagement with content; engaging with content implies attention. Twitter has many different engagement affordances on its platform. Identifying and quantifying these various affordance-driven behaviors across cases of popular and spotlight tweets illustrates which types of behaviors help a user gain attention and interest. Behaviors considered here include:

- *Quote tweet* (also known as retweet with comment): A retweet with comment tweet is an original tweet (not a retweet) where a user is embedding another original tweet in their tweet; they retweet but add their own content as well.
- *Links*: When a tweet has a URL embedded in it.
- *Hashtags*: When a tweet includes at least one hashtag in the text.
- *Location names*: When a tweet includes a formal reference to a location in the tweet text or user profile fields. Note: Users were able to apply a location name of “Hawaii” to their tweet, even if they were not physically based in Hawaii.

6.2.1 Account Proximity and Spotlight Tweets. Through our initial qualitative analysis, we noticed a number of spotlight tweets from users who credentialed themselves as being in Hawaii, either through information in their user bio or through the tweet itself. Examples are seen in Tweets A, B, and C detailed previously. This section further explores this finding through a quantitative analysis across the entire dataset—focusing on features in tweet metadata.

For each user, we examined the user_location, user_bio, and user_utcoffset fields. For both the user_location and user_bio fields, we pragmatically checked for any references to places in Hawaii such as specific islands or cities, towns, and villages (any of the 151 places or eight islands listed on Wikipedia). We marked users who had references to those place or had a user_utcoffset of –36000 (Hawaii’s official and nearly exclusive time zone) as someone tied to Hawaii—what we term *Hawaii-credentialed users*. A limitation of this method is that users who signaled ties to Hawaii through only a tweet itself—for example, visitors on vacation there—were not included.

Table 2. Percentage and Count of Tweets from Hawaii Across Tweet Type Categories

	% in Hawaii	# in Hawaii	Total Number
Spotlight	46.7	99	212
Popular	3.87	16	413
All Other	14.67	14,075	95,972

A total of 46.7% of spotlight tweets were tweets from users who credentialed themselves as being in or from Hawaii, and meanwhile, only 3.87% of popular tweets were from Hawaii-credentialed users (Table 2). A total of 14.67% of all other tweets were from Hawaii-credentialed users. In other words, tweets that were spotlighted were more likely than other tweets to have been posted by a person who lived in or was otherwise associated with the affected area. This is a finding we return to in our discussion in Section 8.

6.2.2 Tweet Features Across Tweet Type. In the following, we examine how different features—including those driven by specific platform affordances like the retweet and quote—are used across our tweet types as well as across users who are or are not connected to Hawaii.

Overall, tweets from users not connected to Hawaii are more likely to be quotes than tweets from users connected to Hawaii. Most notably, the percentage of quoted spotlight tweets from users not connected to Hawaii (18.58%) is nearly double that of spotlight users connected to Hawaii (10.1%). In other words, we see non-locals quoting others more often, suggesting that the non-locals are sharing information with added comments at a higher rate than local users.

We suspect that those who are not connected to Hawaii are quoting those who are in Hawaii, ultimately helping to spread information from locals to a wide network. Overall, spotlight tweets (14.62%) are less likely to be quotes than popular (33.17%) and all other tweets (36.34%). With many spotlight tweets being from Hawaii-credentialed users, this could explain why the overall rate of spotlight tweets being a quote tweet is lower.

6.2.3 Links and Hashtags. Relative to popular (41.16%) and all other (27.9%) tweets, a lower percentage of spotlight tweets (11.32%) include links. A slightly higher percentage of spotlight tweets (16.51%) contained hashtags as compared to popular (10.41%) and all other tweets (12.32%). Both findings are consistent with past work that found positive correlations between retweet count and URLs [87] and hashtags [76].

6.2.4 Location Names. The category with the highest use of location names is spotlight tweets, with 12.26% of these tweets having location names (Table 3). Among the popular tweets and all other tweets categories, users connected to Hawaii use more location names (popular: 0%; all other: 11.36%) than those not connected to Hawaii (popular: 2.76%; all other: 6.86%). Additionally, 10.62% of spotlight tweets from authors not connected to Hawaii contained location names. These results illustrate the deliberate ways local users are drawing attention to themselves by highlighting their proximity to the event.

Taken together, these findings on how different tweet features are used across the different tweet types help explain how spotlight tweets differ from other popular tweets as well as tweets that do not garner as much attention. Spotlight tweets are less likely to be a quote or reply, less likely to contain a link, and more likely to come from Hawaii-credentialed accounts than other popular tweets. All of these distinctions suggest that spotlight tweets were more likely (than other tweet types) to provide original or novel content to the online conversation.

Table 3. Percentage of Quote Tweets and Tweets with Links Within Each Category of Tweets

Type	Quote Tweet (%)			URL (%)			Hashtag (%)			Location Name (%)		
	In HI	Not in HI	All	In HI	Not in HI	All	In HI	Not in HI	All	In HI	Not in HI	All
Spotlight	10.1	18.58	14.62	10.1	12.39	11.32	13.13	19.47	16.51	14.14	10.62	12.26
Popular	35.71	33.08	33.17	14.29	42.11	41.16	7.14	10.53	10.41	0	2.76	2.66
All other	19.22	39.28	36.34	15.82	29.98	27.9	12.71	12.26	12.32	11.36	6.86	7.52

The percentage of quote tweets and tweets with links are higher from users who were not Hawaii-credentialed. Spotlight tweets have lower usage of links. The number of tweets with hashtags is lower for Hawaii-credentialed users. Spotlight tweets have the highest use of location names.

6.3 Tweet Features and Tweet Types: Qualitative Coding and Analysis

Although the tweet metadata offer clues into some of the unique features of spotlight tweets, they do not tell us much about the content of those tweets. Through our qualitative coding, we identified prominent themes and features of tweets about the Hawaii missile crisis as well as some of the unique features of spotlight tweets—compared to popular (other highly retweeted tweets) and other tweets (that were not highly retweeted).

6.3.1 "Rumor" Stance. A first distinction is the "stance" of each tweet toward the imminent threat of an incoming missile. In previous work on rumoring, researchers have differentiated between "spreading" and "correcting" a rumor [42, 74, 88]. We use those same distinctions here, but note that in this event, even those "spreading" the news of an alert were sharing "official" information at the time.

Most tweets about the Hawaii missile crisis were corrections. Only about 11% of the tweets across our samples functioned to spread the news of the missile alert. Of those, about 46% contained some kind of uncertainty about whether or not the threat was real. Interestingly, spotlight tweets—tweets that received unexpectedly high levels of retweets—were more likely than both popular tweets and other tweets to spread the news of the alert without a correction: spotlight (19%), popular (4%), other (10%), $\chi^2(6, N = 572) = 30.75, p < 0.001$ (Table 4). This suggests that spotlighting served, in part, to highlight content from the early moments of the event, before the correction was available.

6.3.2 Informing and Sensemaking. A significant portion of tweets functioned primarily to inform audiences about the threat or the correction (16% of spotlight, 19% of popular, 7% of other tweets) (Table 5); however, other tweets reflected a kind of *sensemaking* activity as their authors attempted to make sense of—or provide frames for others to make sense of—different aspects of the event.

About 15% of spotlight tweets featured some sensemaking about the *threat*—either expressing some uncertainty or actively trying to assess whether or not the alert was real. For example:

> `@corizarek`: "`This Emergency Alert just came to all our phones in Hawaii. No other info coming out and hotel has no shelter. Really hoping it's a false alarm, but if not - send your best vibes this way. #emergencyalert #hawaii <screenshot of alert>`"

> `@davidwolman`: "`I'm in a parking lot in Waimea, Hawaii right now and everyone's phones are buzzing with a warning about an incoming missile threat and "This is not a drill" language. Info? Anyone?`"

The preceding tweets, posted minutes after the initial alert, underscore the fear and uncertainty that many people in Hawaii experienced that morning. Both were spotlight tweets. Popular tweets (5%) and other tweets (9%) were significantly less likely to include that kind of deliberation.

Table 4. Percentage of Tweets by Tweet Type and Stance

	Spotlight Tweets (%)	Popular Tweets (%)	Other Tweets (%)
Spreading	18.91	4.1	9.77
Correcting	77.11	88.21	77.59
Neutral	3.98	7.69	12.64

Table 5. Proportion of Tweets in Each Group Coded in Each Sensemaking Type

	Spotlight Tweets (%)	Popular Tweets (%)	Other Tweets (%)
Informing	15.92	18.97	13.22
Sensemaking About the Threat	14.93	5.13	8.62
Sensemaking About Protective Actions	6.47	1.03	4.6
Sensemaking About Causes, Impacts, Solutions	28.86	41.03	32.18
Sensemaking About Responses	13.43	15.38	10.92
Other	20.4	18.46	30.5

A small but meaningful number of spotlight tweets noted *protective actions*—including tweets asking for advice about what protective actions to take, tweets explaining the protective actions that people were actually taking, and tweets (often after the correction) lamenting the lack of preparedness and inability for people to take productive protective measures.

The most prominent category of tweets was sensemaking about *causes, impacts, and solutions* (34% of all the tweets we coded, and 29% of spotlight tweets). Tweets in this category attempted to determine (or to guide others in determining) why the false alarm had occurred, why it had been so plausible and therefore so frightening, and what should be done in the future to prevent a similar event in the future. For example, the following popular tweet was a follow-up posted by @brianschatz (author of Tweet D, presented earlier):

> `@brianschatz:` "`AGAIN FALSE ALARM. What happened today is totally inexcusable. The whole state was terrified. There needs to be tough and quick accountability and a fixed process.`"

Many tweeters assigned blame to the employee who "pushed the wrong button" and demanded that someone be held accountable. Others speculated that the design of the technology or the process had led to the "human error." Along another trajectory, several tweeters cited the underlying geopolitical conditions (relations between North Korea and the United States) or the always-looming threat of nuclear war as contributing to the panic. Interestingly, sensemaking about causes, impacts, and solutions was somewhat less common in spotlight tweets (29%) than in popular tweets (41%) and other tweets (32%), $\chi^2(14, N = 572) = 34.76, p < 0.01$. Highly retweeted tweets from accounts that were already relatively visible (in other words, existing influencers) were more likely to participate in this event by speculating about its causes and assigning blame.

A related category of sensemaking was around various responses to the false alarm—what actions people and organizations took or should have taken to address the false alarm. Many of these tweets were critical of local authorities for waiting too long to send out an official correction.

> `@TreyYingst:` "`It took nearly 40 minutes for residents of Hawaii to receive a second alert informing them that a ballistic missile was not incoming. 40 minutes.`"

A significant number of sensemaking tweets—especially those around responses and causes, impacts, and solutions—included a "target" where the author located responsibility or blame (and in a few cases praise or credit). For these tweets, we coded (1) whether the tweet was criticizing or defending, (2) who the tweet was criticizing/defending, and (3) what the tweet was criticizing/defending them for. We found that whereas popular tweets were more likely to be complaining about the false alarm (25% spotlight, 31% popular, 24% other), spotlight tweets are more likely (compared to other tweet types) to be critical about timing and communication of the correction (13% spotlight, 6% popular, 5% other), $\chi^2(14, N = 572) = 26.36, p = 0.02$.

"Who" these tweets were targeting also varies. Spotlight tweets were more likely to blame local authorities (14% spotlight, 6% popular, 6% other) or the specific employee (7% spotlight, 4% popular, 7% other)—often demanding that "someone be fired." Popular tweets, however, were more likely to politicize the crisis (12% spotlight, 23% popular, 17% other (Chi-squared: 7.76; p-value: 0.10; DOF: 4)) and blame a political figure—especially then-President Donald Trump (5% spotlight, 12% popular, 10% other), $\chi^2(48, N = 572) = 67.9, p = 0.03$.

> @Anonymized_User: "When is Trump going to take a break from golfing to address Hawaii's fake missile threat?"

> @bobcesca_go: "Let's be clear about this: Hawaii wouldn't be on nuclear alert in the first place were it not for Trump and his spastic recklessness. <quoted tweet>"

Among tweets politicizing the crisis, many criticized Trump for either his responses to the crisis (e.g., not making an announcement because he was on the golf course) or contributing to conditions of political instability that created the need for the test that morning and/or made the threat plausible and especially frightening.

6.4 Spotlight Tweets Featured Proximity Signals

Another salient aspect of spotlight tweets involved signals, within the tweet text, of "proximity" to the event. In some cases, the tweet text indicated—either explicitly or implicitly—that the tweeter was in Hawaii, positioning them as an eyewitness to the event. In other cases, the tweets indicated a kind of "emotional proximity" [25] to the event, including a personal connection to someone there, a "place attachment" [23] to Hawaii, or a sense of connection due to the tweet author experiencing a similar kind of event in the past. Aligned with our quantitative analyses of tweet features previously, our systematic qualitative analysis revealed spotlight tweets (57%) to more often include indications of proximity (either physical or emotional) than popular (14%) and other tweets (24%), $\chi^2(12, N = 572) = 122.66, p < 0.001$. This aligns with our other findings, across the broader dataset (not just the hand-coded tweets), that spotlight tweets were more likely to be posted by people whose metadata (time zone) indicated they were in Hawaii.

6.4.1 Spotlight Tweets Shared Personal Accounts of the Missile Threat. Like Tweet C presented earlier, spotlight tweets were also more likely (than popular and other tweets) to feature personal accounts of the alert—in other words, how people experienced receiving the alert and the time period between the alert and its correction. For example:

> @brynguist: "At 8:07am everyone in Hawaii got a phone alert: BALLISTIC THREAT INBOUND TO HAWAII. SEEK IMMEDIATE SHELTER. THIS IS NOT A DRILL. The next 10 minutes were the most terrifying of my life, until I finally checked twitter and saw this. But seriously, WTF just happened <quoted tweet: @Hawaii_EMA: NO missile threat to Hawaii.>"

> @Anonymized_User: "A cop pulled up to my family's home back in Hawaii. She stepped out of her car, crying. Told my family it was a mistake, that there wasn't a missile coming... THEY HAD PEOPLE IN HAWAII BELIEVING TODAY WAS THEIR LAST DAY TO LIVE."

Table 6. Percentage of Personal Accounts by Tweet Type

	Spotlight Tweets (%)	Popular Tweets (%)	Other Tweets (%)
Personal Accounts	49.75	11.28	10.34
First-Hand Personal Accounts	41.29	6.67	5.75
Second-Hand Personal Accounts	8.46	4.62	4.6

Table 7. Follower Changes by Time Across Different Tweet Authors

	Mean (%)	90th Percentile (%)	Standard Deviation (%)
Spotlight	1.12	0.75	5.31
Popular	0.03	0.04	0.16
All Others	0.12	0.06	7.55

The preceding tweets are just 2 of the 124 (22%) tweets we coded with often heart-wrenching personal accounts of the fear and uncertainty that people experienced between the alert and its correction. As Table 6 shows, 50% of spotlight tweets included personal accounts (many of them first-hand accounts)—significantly more than popular (11%) and other (10%) tweets, $\chi^2(6, N = 572) = 115.41, p < 0.001$.

Another interesting feature of spotlight tweets—and to some extent popular and other tweets as well—was the inclusion of a screenshot of the text message alert (or official correction) that either the tweet author (first-hand), their close relation (second-hand), or another person (third-hand or other) had received.

Spotlight tweets (15%) were more likely to contain embedded screenshots of the alert or its correction than popular tweets (7%) and other tweets (5%), $\chi^2(4, N = 572) = 14.1, p < 0.01$. And when tweets did contain a screenshot, those were more often first-hand shots (of their own phone) in spotlight tweets (87%) than in popular tweets (31%) and other tweets (50%).

6.5 Follower Gains and Spotlight Tweets

We also observed the effect of spotlight tweets on the user's online social position. Within our dataset, if a user tweets out a spotlight or popular tweet followed by at least one other tweet, then we can calculate their follower gain during that period of time. For users who have neither a spotlight or popular tweet, we calculate their change in followers as the max(user followers)–min(userfollowers) within those 2 hours.

Our findings illustrate that spotlight tweet authors have the highest mean change in followers; however, this could be driven by a few users given the high standard deviation (Table 7). Popular tweet authors experienced lower follower growth compared to all others, and this might be due to the fact that popular authors already have a large following—resulting in smaller percentage changes in follower count. In aggregate, this shows that being a spotlight tweet author is accompanied by increases in followers in a way that is different from being a popular tweet author.

One user's follower growth in particular was of note. House Representative Tulsi Gabbard's spotlight-and-popular tweet, sent on January 13, 2018, at 18:19:30, definitively stated that the alarm was false. Gabbard then retweeted Hawaii EMA's message of "`RT @Hawaii_EMA: NO missile threat to Hawaii`" on the same date at 18:28:37. During this 9-minute window, Gabbard gained 710 followers, going from 167,185 to 167,895 followers (a 0.42% increase in followers). This rapid gain of followers illustrates that spotlight tweets could draw not only temporary but also potentially longer lasting attention to the user.

7 GENERALIZABILITY OF THE SPOTLIGHT TWEET

Although our work focuses on spotlight tweets within a single event (the Hawaii missile crisis), the phenomenon of spotlight tweets can be found in many other high-attention events. In the following, we provide examples of spotlight tweets within the COVID-19 pandemic and the 2020 presidential elections.

One prominent example from the COVID-19 pandemic is the case of Dr. Eric Ding (@DrEricDing). On January 25, 2020, the self-described "public health scientist" posted an attention-grabbing tweet about the novel coronavirus:

> "HOLY MOTHER OF GOD - the new coronavirus is a 3.8!!! How bad is that reproductive R0 value? It is thermonuclear pandemic level bad - never seen an actual virality coefficient outside of Twitter in my entire career. I'm not exaggerating... #WuhanCoronovirus #CoronavirusOutbreak"

At the time, Ding's account had only 2,131 followers, and prior to his engagement with the coronavirus discussion, his tweets received on average less than one retweet per tweet. But this tweet, which became the root of a long thread, received 15K retweets, several orders of magnitude more than would otherwise be expected according to his baseline retweet rate (see Figure 12 in appendix). The tweet was retweeted and quoted by several high-follower accounts, including journalists and social media influencers. In less than a day, @DrEricDing's account accumulated nearly 30K followers. Ding soon updated his Twitter profile to include the title of "epidemiologist," perhaps to reflect the informational role that he had begun to assume. And despite criticism for his inflammatory language, he quickly rose—on top of these initial attention gains—to become a leading voice in the public discussion of what would soon be known as COVID-19 [9].

Ding's tweet represents a particularly salient example of a spotlight tweet from the COVID-19 discourse; however, we have also seen the phenomenon during the 2020 U.S. election, where spotlighting appears to have played a role in highlighting perceived voting issues, taking local rumors into national conversation, and pulling them into a false narrative of massive voter fraud—in what has become known as "the Big Lie." For example, spotlight tweets were integral to the development and spread of #SharpieGate, which grew from a collective sensemaking process around Sharpie pens bleeding through ballots, to a false conspiracy theory that the pens were intentionally given to conservative voters to invalidate their votes, as part of a voter fraud scheme to cause President Trump to lose the election [17]. As #SharpieGate began to gain steam early in the morning on the day after the election (November 4, 2020), a number of spotlight tweets—many of them from voters in Arizona who had sincere concerns about their ballots—gained outsized attention from the crowd. A few of these spotlight tweets were extremely highly retweeted (e.g., an account with 291 followers that received 8,941 retweets), often due to boosts by large-following accounts, including accounts of political figures and partisan media, which helped them "go viral." Although retweets of spotlight tweets (shown in orange in Figure 11) and popular-not-spotlight tweets (shown in blue) both contributed to the early growth of #SharpieGate (4:00–14:00 UTC), as the conspiracy theory took off later that day (17:00–20:00 UTC), attention shifted onto more "popular" content provided by political figures and media pundits, who pushed the story out to larger, national audiences. See Figure 13 in the appendix for a more detailed view of these dynamics.

From continued research on the use of Twitter during crises and other breaking news events, we perceive that spotlight tweets are a common feature of high-attention events, and that the concept can be used as a lens for understanding these events in ways that other approaches may miss.

7.1 Generalizing the Definition of Spotlight Tweets to Spotlight Social Media Posts

As events range in duration and popularity, the metrics used to calculate what is considered to be a spotlight tweet can vary. The parameters used here are unique to our study and will most likely

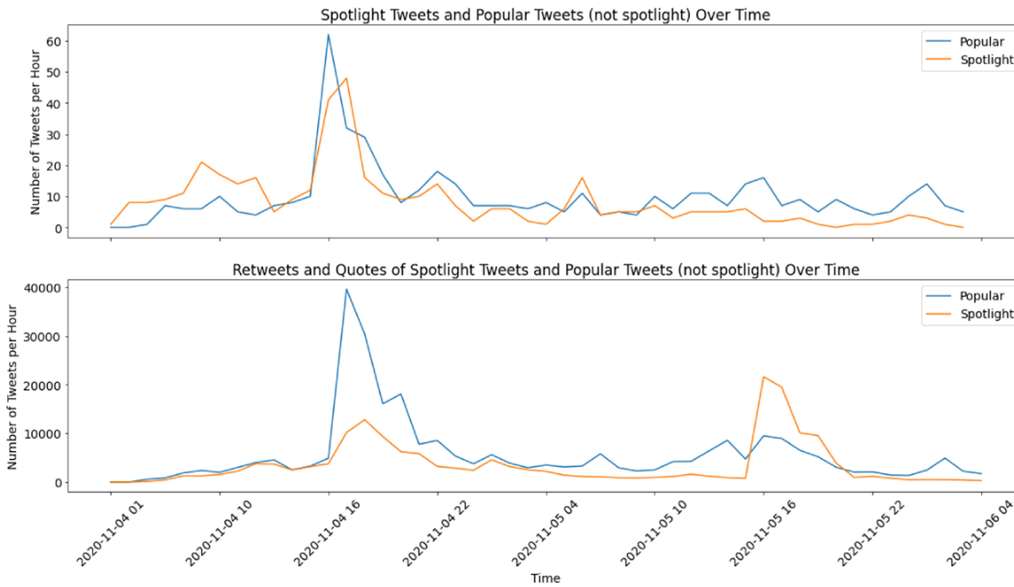

Fig. 11. Top: Number of spotlight tweets and popular-not-spotlight tweets over time. Bottom: Number of retweets and quotes of spotlight tweet and popular-not-spotlight over time.

need to be adjusted across events as appropriate. The concept of a spotlight social media post likely also exists on other platforms where users propagate another user's content toward their own audience. However, algorithms of how user-shared content is promoted and displayed will differ, which could dampen or heighten the impact of spotlighting. Explorations of how spotlighting occurs and impacts attention dynamics across platforms is an exciting future direction for this work. The high-level takeaway is that spotlight social media posts are identified by comparing a post's attention to a user's typical baseline rather than simply looking at raw number of shares. By doing so, we are selecting social media posts that have received an unexpectedly large amount of attention—highlighting cases where surprising and potentially meaningful shifts in attention have occurred. Spotlight social media posts provide a unique window for studying attention dynamics, and comparing them to non-spotlighted social media posts provides additional insight into why some information "goes viral" during crises and other emerging events.

8 DISCUSSION

Our results highlight the various characteristics of spotlight tweets, including the features (fewer links, more location names, less likely to be quote tweets) that distinguish them from popular tweets. Through our analysis, we also gained new insight into how information flows online, including the role of spotlight tweets—and spotlighting—in shaping how attention is conferred in both the short and long term. Next, we discuss some of the most salient themes and implications of this research.

8.1 Visualizations for Attention Dynamics

Our study highlights the value of visual, exploratory methods in social media analysis. We use a cumulative retweet graph that reveals possible mechanisms for how retweets and their downstream exposure impact the overall spread of the information—which is reflected in both the shape of the signature and the placement and size of specific accounts. The size of data points in a cumulative graph can be used to represent additional features—in this case, the follower count of the sharing user. Incorporating user follower counts in this way highlights the users whose audiences likely contributed to the spread of spotlight tweets.

This visualization technique allows us to examine the attention dynamics on social media in depth. Applying a bubble graph technique to a cumulative graph, we overlay a signal of audience size onto a temporal graph—so that we can see the role of high visibility accounts in shaping information trajectories. This technique is a tool to support this kind of mixed-method inquiry into Twitter discourse—and particularly for understanding the impacts of exposure on propagation.

A typical pattern that we observed while exploring this visualization for spotlight tweets was (1) a low retweet rate in the early stages of a tweet followed by (2) retweets from one or more large-follower accounts, typically resulting in (3) a higher retweet rate in the subsequent minutes. Visualizations, such as those used in this article, lend themselves to mapping the complex space of engagement mechanisms and information dynamics, which requires looking at information on different measures and possible mechanisms at the same time. Being able to analyze the growth of a tweet and who engages with that content, rather than just its content and final retweet count, enables us to understand the possible mechanisms associated with a tweet's growth trajectory. Although these techniques do not allow for causal claims, they are valuable for gaining insight into factors that likely play a role in structuring information flow in these systems.

8.2 Spotlight Operators and the Role of Journalists in Conferring Attention and Credibility

Our multi-modal visualization allowed us to investigate an underlying dynamic of many spotlight tweets in this study: the process of a user with a high-follower count conferring their attention onto a spotlight tweet and its author. We call this behavior *spotlighting*. Spotlighting functions to give the tweet author access to a "new" audience (that of the spotlight operators) and could potentially lead to a persistent increase in the spotlight tweet author's audience. This effect appears in the distinctive S-shape signature in the cumulative retweet visualization, where inflections in this curve coincide with retweets by users with large audiences.

In our data, there were several examples of self-described journalists spotlighting others—focusing attention onto a local user, catalyzing the spread of a tweet, and serving as a bridge between one community or audience and another. Previous work has described the role of journalists as gatekeepers in news [30, 63, 64] and online [5], whereas others have noted how this role has been disrupted by information-sharing dynamics on the internet [6, 19]. In their role as spotlighters, journalists are not quite gatekeepers or gatewatchers [6], but they are serving a hybrid role of identification, selection, and amplification—guiding the attention of their audiences towards certain content and specific authors. For example, for Tweet B, a digital producer at Vanity Fair shines the spotlight on the tweet's author by retweeting and spreading Tweet B to their audience, which included a number of technology enthusiasts. In Tweet C, retweets from one journalist to another helped start a cascade of retweets. These findings are consistent with work from Kwak et al. [39], as well as Jürgens et al. [33] and Wei et al. [81], showing how journalists still play a significant role in shaping information flows, even through "social" media.

Tweets that are spotlighted then appear on another user's radar most likely not because that user made an explicit decision to engage with the spotlight author, but rather because a journalist or other influencer directed attention to it. In contrast to the growth of a popular tweet (Tweet D), where retweets are immediate and most likely from users who are already attending to the author, spotlight tweets need time and help from others to gain that attention. However, once an initial volume of attention is placed on the user, it can then become easier to gain more and more attention (resulting in an S-shape pattern of retweet growth), congruent with a "rich get richer effect" (which is termed the *Matthew effect* [44]). Although the examples discussed here emerged through our qualitative analysis of a sample of spotlight tweets for this specific event, we plan to more systematically explore the role of journalists as facilitating visibility of local information and local accounts during times of crisis in future work.

Looking at the intersection of attention dynamics and crisis events illustrates how specific voices can be amplified, both by their own credentials and the work of others, such as journalists. It also highlights how this amplification has the potential to impact collective sensemaking: amplified voices can become key figures and sources of information that is being processed and synthesized. It is important to recognize that spotlighting can, theoretically, lead to drastic shifts in sensemaking as information is moved around the social network in particular ways. For example, actions based on the (unofficial) confirmation of a false alarm, present in Tweet B, could have been taken sooner for those monitoring social media than by those waiting for official guidance—and indeed, our systematic qualitative analysis revealed several tweets where people remarked and even complained that social media users became aware of the correction long before others in Hawaii. In this case, through a process of spotlighting and sensemaking, the online crowd had a more up-to-date understanding of the event than those only receiving information from more traditional broadcast channels.

8.2.1 Advantages of Spotlighting: Local Credibility as Salient Information. One key benefit of spotlighting is the ability to elevate the voices of those who are directly impacted by the crisis event. Our research suggests that one reason people and their posts gain attention during a crisis event is related to their newfound position as credible or relevant information providers due to their proximity to the event. In our metadata analysis, 46.7% of spotlight tweets were authored by users connected to Hawaii compared to 14.68% of tweets overall. Similarly, through our qualitative coding, we found that spotlight tweets were more likely (than other tweets) to mark their authors as being in Hawaii. Being connected to the physical location of the crisis event likely increases a user's perceived credibility as well as the potential for novelty in their content; both the user and their posts may see increased attention as a result. This is consistent with previous research [70, 84] demonstrating that locals who are on the ground of the crisis gain visibility through retweets (often by non-locals).

In our case study, local credibility is also paired with a unique experience of the event: the people in Hawaii were the only people to receive an official alert and to have first-hand experiences of its immediate aftermath. Local tweeters, as eyewitnesses, were sources of novel, significant, and, in many cases, emotionally laden information that other users found valuable. Whether or not spotlight tweet authors connected to Hawaii intended to draw attention to themselves, the effect was that a high percentage of spotlight tweets came from users connected to Hawaii who were sharing screenshots of the alert and personal accounts of its aftermath. Spotlighting therefore acts as the bridge from locals to a global audience, bringing relevant information out of the local context to a broader community as non-locals try to understand and make sense of the event from afar. Again, this could be a vital component for effective and efficient collective sensemaking. Often it is those who are not directly affected by crisis events who have the capacity to bring together information from disparate sources into a more complete picture of dynamic crisis response activities [12, 72]. Spotlighting can aid in this process, bringing attention to important information.

8.2.2 Challenges of Spotlighting: Rapid and Potentially Unchecked Spread of Information. Although spotlighting has the potential to shed light on more obscure pieces of information, the process has few checks and balances in place, allowing information to rapidly gain attention and spread through the social network through a chain of relatively few individual decisions to share the content; with each of these decisions, the message may move farther away from its original author and original context. In our findings, we highlight that Tweet B (an early confirmation that this was a false alert), posted by a user with 160 followers, received 1,071 retweets (10.63:1 retweets to followers), whereas a tweet carrying a similar message from the official ABC network Twitter account, which had 13,259,900 followers, received 1,524 retweets (0.0001:1 retweets to followers). Despite beginning with a much smaller audience of followers, the combination of timing, content, author location, and the underlying network structure of Twitter at the time resulted in Tweet B receiving a comparable number of retweets to a tweet posted by a national news network. Repeatedly we see users with a small following gain the same amount of attention and engagement with their content as an established news network. Social media platforms provide the conditions that allow for spotlight posts to occur, removing the barriers to distribution [3] and giving users of all audience sizes the potential to gain exposure and quickly spread information to a large public. If attention is a zero-sum game, then this increase in attending to typically low-visibility accounts means a decrease in attention somewhere else. This has implications on the way sensemaking unfolds during crisis events. The interplay between existing influencers and newly relevant accounts in sensemaking during crisis events is something that warrants more exploration.

Spotlight tweets can shed light on critical information that would normally not reach wider audiences. Spotlighting diverts the attention of audiences to focus on content and users being retweeted by high-follower accounts, and perhaps implicitly defers attention from others. The effect of this shift in attention can be dramatic both in its reach and in the social position of the spotlighted user immediately afterward. Although this spotlighting can help surface important (and novel) information, it can also function to highlight and (potentially) further propagate misinformation and disinformation. Even when it appears that users have credibility, either through explicitly credentialing themselves as a local or through being retweeted from high-status accounts or journalists, the tweets and the information they contain are not necessarily officially verified.

8.3 Long-Term Impact of Spotlight Tweets

Last, the impact of spotlighting may not be isolated to just an increase in attention toward a single tweet in time. In this article, we begin to explore possible long-term social dynamics of spotlighting.

Borrowing from work on technology within organizations [51, 52], it can be productive to view social media as sociotechnical systems, where actions taken by users shape the underlying structure of the environment, which in turn shapes future actions [34, 57, 68]. This perspective suggests that spotlight incidents can have lasting effects on the underlying structure of the social network, such as creating new and/or reinforcing existing pathways for information exchange, lifting the status of particular accounts, and triggering new recommendation mechanisms. Previous empirical research showing how attentional convergence during crisis events onto the accounts of official responders leads to follower growth supports this view [67].

Additionally, some social media users have recognized that convergence during crisis events can be an attention opportunity and have developed strategies to exploit these situations. For example, social media "influencer" Mike Cernovich has boasted about using crisis events to grow his audience [43]. However, our analyses do not support a conclusion that during the Hawaii event a substantial proportion of spotlight tweeters (or spotlight operators) were simply seeking attention, but rather that they were primarily motivated by sharing relevant information. Interestingly, in this case, it was popular tweets—highly retweeted tweets from users who already had significant followings—that functioned (more often) to politicize the event. This supports the possibility that some portion of influential accounts have built their followings—or at least maintain their attention—by quickly converging onto the virtual scene of a crisis event and politicizing it for audiences that receptive to polarizing content.

9 IMPLICATIONS FOR RESEARCH AND DESIGN

Correctly Highlighting Locals and Experts During Crisis Events. During a crisis event, the public turns to social media for event critical information; however, there often exists an abundance of information, and it can be hard for users to decide where to focus their attention. At the same time, there are certain individuals who can suddenly gain a spotlight on their account, often by being promoted by other high-profile users. During these information-critical times, it can be valuable to provide visibility to information provided by people with relevant expertise, whether they be a local resident or a relevant expert (e.g., an epidemiologist or seismologist). This sort of signaling is already happening through the behaviors of spotlight operators, but designers should investigate how to support this type of spotlighting without enabling those who may want to exploit the crisis event for political or reputational gain. On Twitter, there exists a universal verification (blue) checkmark that indicates to platforms users that an account of public interest is authentic. We encourage designers to investigate the safety benefits of highlighting verified locals and experts during crisis events. To some extent, this kind of policy is already happening on Twitter in the context of COVID-19, as the platform has worked to quickly verify accounts with relevant expertise [41, 77]. We also see this happening with the 2020 U.S. elections as Twitter aims to add identification markers on political candidates' accounts.[12] This sort of official signaling across crisis events can help not only users but also spotlight operators who are seeking information to promote to their audiences.

Building Visualization Tools to Help Users Critically Evaluate Information Trajectories. A persistent challenge for social media users is trying to assess the quality and veracity of information they encounter. One dimension of this issue is that although users can see overall engagement numbers and can typically find the original post in an information cascade, they are not able to see how a post reaches them. Insight into the information trajectory of a piece of content could help users make more informed decisions about that content—not just about its source but about why it is spreading. This may be particularly relevant for those who risk becoming unwitting agents in the spread of disinformation—that is, false or misleading content that is seeded and/or spread for a strategic objective [68]. For example, seeing that a politically motivated "influencer" account helped catalyze the spread of a piece of content could motivate downstream users to be thoughtful about whether and how they engage. Researchers have argued that building additional signals about information into platform design could support media literacy [29]. The visualization technique used here allowed us, as researchers, to identify these otherwise hidden features, and this approach could be further developed into a visualization tool to support social media users in critically analyzing the information in their feeds.

Identifying Long-Term Structural Change from Spotlight Operators. Our findings demonstrate that there are likely long-term effects that result from the behavior of spotlight operators. In the short term, a spotlight is placed on an account, but looking in the long term we see that there are also lasting gains in follower accounts for spotlight tweet authors. Their social position within the platforms could be forever changed as a result of their participation in event-related conversations. Although investigating the impact of long-term structural changes in the underlying social network is outside the scope of this work, we encourage social media platform designers to develop metrics and track the impact (both short and long term) of spotlight operators. In particular, we believe that it will be important to understand how this spotlighting behavior and resulting structural change impacts the flow of information online. Gathering longitudinal observations of this phenomena could help identify how and why social media platforms grow and evolve over time and, if necessary, identify any malicious or coordinated behaviors.

[12] About election labels on Twitter: https://help.twitter.com/en/using-twitter/election-labels.

10 LIMITATIONS

The presented study focuses on one event with a relatively short time span. Additional case studies that vary in type and location will deepen our understanding of the spotlighting phenomenon. Our study focuses on a particular crisis event, but we suspect that spotlight tweets also appear in other contexts and domains. As explained in detail in Section 5, our data collection over-samples tweets from accounts with affiliations in Hawaii—which likely means that the differences between "spotlight" and "other" tweets on dimensions related to proximity are under-stated here (compared to a more random sampling of tweets about the crisis). Our collection also has the potential bias of missing spotlight tweets from high-activity users due to the rate limits of the Twitter API in pulling historical activity of users; we are unable to establish a pre-event baseline for attention for highly active users. To address this limitation, we explored two additional methods of assessing historical averages and found that our findings on correlations between tweet features and spotlight tweets still hold. Additionally, our data might be missing user-deleted tweets due to the timing of the data collection, which took place after the event.

11 FUTURE WORK

We identify several directions for future research. The first includes understanding patterns of spotlight posts longitudinally, including the long-term effects on a spotlight tweet author's social position. The second includes conducting user interviews to better understand the imagined audiences, motivations, and intentions of spotlighted users and spotlight operators. The third is that comparing the behaviors of spotlight tweets across multiple events can help us better understand how our findings generalize. We also believe that designing experiments that can directly measure the impact of spotlight tweets on the sensemaking process can be informative and yield new insights for design. Last, we believe that the concept of a spotlight post can extend beyond Twitter and crisis events, and we are curious to see how this phenomenon generalizes to other social media platforms and other contexts.

12 CONCLUSION

In this empirical case study, we introduced the concept of a social media spotlight tweet, a tweet that receives an unexpected amount of attention. We investigated how the tweet's feature, the author's profile, the network structures, and the social dynamics of Twitter each play a role in the way spotlight tweets are formed and grown. We demonstrated that spotlight tweets offer a productive window into the outsized role of users who lack existing influence and discussed the impact spotlight tweets have on collective sensemaking. Last, we used a visualization technique that enables us to identify and explore the role of spotlight operators. Future work should examine spotlight posts for other events and contexts to determine the robustness of these behavioral findings and how well they generalize to different domains.

APPENDIX

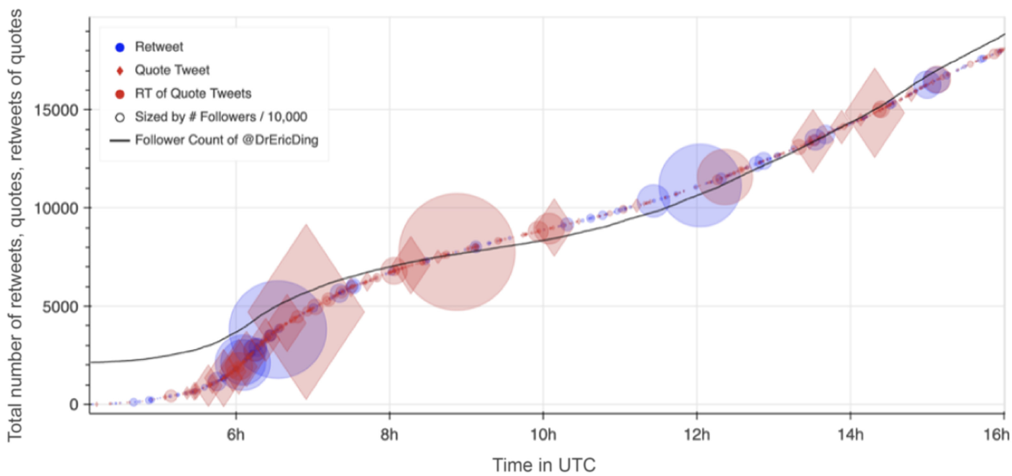

Fig. 12. Cumulative engagement graph of @DrEricDing's "MOTHER OF GOD" tweet. This graph plots time on the x-axis and the cumulative number of engagements (retweets, quote tweets, and retweets or quote tweets) of @DrEricDing's tweet on the y-axis. Individual tweets are plotted, sized, and shaped by tweet type and relative to follower size (divided by 10,000 for absolute size). This graph is a more recent iteration on the visualization technique explored in depth in this article.

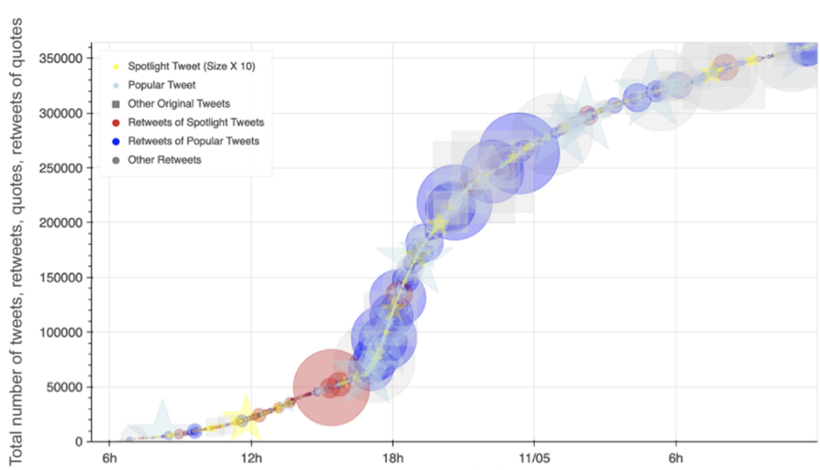

Fig. 13. Cumulative tweet graph for #SharpieGate. This graph plots time on the x-axis and the cumulative number of tweets (including originals, retweets, quote tweets, and retweets or quote tweets) related to #SharpieGate on the y-axis. Individual tweets are plotted, sized, and shaped by tweet type and relative to follower size (divided by 20,000 for absolute size). Spotlight tweets (that received more retweets than expected) are yellow-colored stars (and enlarged by a multiplier of 10). Popular tweets (that received many retweets but not an unexpected amount for the account) are light blue stars. The graph shows a marked shift from almost exclusively spotlight tweets (and retweets of those tweets) to a much larger proportion of popular tweets around 18:00 UTC.

ACKNOWLEDGMENTS

We thank John Robinson for building the data infrastructure used to collect tweets for this analysis, and Pushshift for providing additional data used to confirm our determinations of spotlight and popular accounts. We also thank the reviewers for their thoughtful commentary and peers who have offered suggestions and feedback.

REFERENCES

[1] Ahmer Arif, Kelley Shanahan, Fang-Ju Chou, Yoanna Dosouto, Kate Starbird, and Emma S. Spiro. 2016. How information snowballs: Exploring the role of exposure in online rumor propagation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'16)*. ACM, New York, NY, 466–477. https://doi.org/10.1145/2818048.2819964

[2] Eytan Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J. Watts. 2011. Everyone's an influencer: Quantifying influence on Twitter. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM'11)*. ACM, New York, NY, 65–74. https://doi.org/10.1145/1935826.1935845

[3] Danah Boyd. 2010. Streams of content, limited attention: The flow of information through social media. *EDUCAUSE Review* 45, 5 (2010), 26.

[4] Danah Boyd, Scott Golder, and Gilad Lotan. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*. IEEE, Los Alamitos, CA, 1–10. https://doi.org/10.1109/HICSS.2010.412

[5] Peter Bro and Filip Wallberg. 2014. Digital gatekeeping: News media versus social media. *Digital Journalism* 2, 3 (2014), 446–454.

[6] Axel Bruns. 2003. Gatewatching, not gatekeeping: Collaborative online news. *Media International Australia Incorporating Culture and Policy* 107, 1 (2003), 31–44.

[7] Axel Bruns, Jean E. Burgess, Kate Crawford, and Frances Shaw. 2012. *#qldfloods and @QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods*. Research Report. ARC Centre of Excellence for Creative Industries & Innovations (CCI).

[8] Deborah Bunker, Milad Mirbabaie, and Stefan Stieglitz. 2017. Convergence behaviour of bystanders: An analysis of 2016 Munich shooting Twitter crisis communication. In *Proceedings of the 16th IEEE/ACIS International Conference on Software Engineering (ACIS'17)*.

[9] Ellen Burstein and Kendall Shields. 2021. The Complicated Prophet of the Pandemic. Retrieved January 24, 2023 from https://www.thecrimson.com/article/2021/3/4/feigl-ding-convo/.

[10] Kathy Charmaz and Liska Belgrave. 2012. Qualitative interviewing and grounded theory analysis. In *The SAGE Handbook of Interview Research: The Complexity of the Craft* (2nd ed.). SAGE Publications, 347–365.

[11] Louise K. Comfort, Kilkon Ko, and Adam Zagorecki. 2004. Coordination in rapidly evolving disaster response systems: The role of information. *American Behavioral Scientist* 48, 3 (2004), 295–313.

[12] Dharma Dailey and Kate Starbird. 2014. Journalists as crowdsourcerers: Responding to crisis by reporting with a crowd. *Computer Supported Cooperative Work* 23, 4-6 (2014), 445–481.

[13] Elliott R. Danzig, Paul W. Thayer, and Lila R. Galanter. 1958. *The Effects of a Threatening Rumor on a Disaster-Stricken Community*. National Academies Press, Washington, DC.

[14] Sarah Elsharkawy, Ghada Hassan, Tarek Nabhan, and Mohamed Roushdy. 2016. Towards feature selection for cascade growth prediction on Twitter. In *Proceedings of the 10th International Conference on Informatics and Systems*. 166–172.

[15] Casey Fiesler and Nicholas Proferes. 2018. "Participant" perceptions of Twitter research ethics. *Social Media+ Society* 4, 1 (2018), 2056305118763366.

[16] Megan Finn. 2018. *Documenting Aftermath: Information Infrastructures in the Wake of Disasters*. MIT Press, Cambridge, MA.

[17] Center for an Informed Public, Digital Forensic Research Lab, Graphika, and Stanford Internet Observatory. 2021. *The Long Fuse: Misinformation and the 2020 Election*. 1.2.0. Election Integrity Partnership. https://stacks.stanford.edu/file/druid:tr171zs0069/EIP-Final-Report.pdf.

[18] Charles E. Fritz and John H. Mathewson. 1957. *Convergence Behavior in Disasters: A Problem in Social Control*. National Research Council of the National Academy of Sciences, Washington, DC.

[19] Dan Gillmor. 2006. *We the Media: Grassroots Journalism by the People, for the People*. O'Reilly Media.

[20] Vincent F. Hendricks and Mads Vestergaard. 2019. *Reality Lost: Markets of Attention, Misinformation and Manipulation*. Springer Nature.

[21] Thomas Heverin and Lisl Zach. 2010. *Microblogging for Crisis Communication: Examination of Twitter Use in Response to a 2009 Violent Crisis in the Seattle-Tacoma, Washington, Area*. ISCRAM.

[22] Thomas Heverin and Lisl Zach. 2012. Use of microblogging for collective sense-making during violent crises: A study of three campus shootings. *Journal of the American Society for Information Science and Technology* 63, 1 (2012), 34–47.

[23] M. Carmen Hidalgo and Bernardo Hernandez. 2001. Place attachment: Conceptual and empirical questions. *Journal of Environmental Psychology* 21, 3 (2001), 273–281.

[24] Starr Roxanne Hiltz and Linda Plotnick. 2013. Dealing with information overload when using social media for emergency management: Emerging solutions. In *Proceedings of the 10th International Conference on Information Systems for Crisis Response and Management (ISCRAM'13)*.

[25] Y. Linlin Huang, Kate Starbird, Mania Orand, Stephanie A. Stanek, and Heather T. Pedersen. 2015. Connected through crisis: Emotional proximity and the spread of misinformation online. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'15)*. ACM, New York, NY, 969–980.

[26] Amanda Lee Hughes and Leysia Palen. 2009. Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management* 6, 3-4 (2009), 248–260.

[27] Amanda L. Hughes, Leysia Palen, Jeannette Sutton, Sophia B. Liu, and Sarah Vieweg. 2008. Site-seeing in disaster: An examination of on-line social convergence. In *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM'08)*. 44–54.

[28] Cindy Hui, Yulia Tyshchuk, William A. Wallace, Malik Magdon-Ismail, and Mark Goldberg. 2012. Information cascades in social media in response to a crisis: A preliminary model and a case study. In *Proceedings of the 21st International Conference on World Wide Web (WWW'12)*. 653–656.

[29] The Aspen Institute. 2021. *Commission on Information Disorder Final Report*. Aspen Digital. https://www.aspeninstitute.org/publications/commission-on-information-disorder-final-report/.

[30] Morris Janowitz. 1975. Professional models in journalism: The gatekeeper and the advocate. *Journalism Quarterly* 52, 4 (1975), 618–626.

[31] Maximilian Jenders, Gjergji Kasneci, and Felix Naumann. 2013. Analyzing and predicting viral tweets. In *Proceedings of the 22nd International Conference on World Wide Web (WWW'13)*. ACM, New York, NY, 657–664.

[32] Hawoong Jeong, Zoltan Néda, and Albert-László Barabási. 2003. Measuring preferential attachment in evolving networks. *Europhysics Letters* 61, 4 (2003), 567.

[33] Pascal Jürgens, Andreas Jungherr, and Harald Schoen. 2011. Small worlds with a difference: New gatekeepers and the filtering of political information on Twitter. In *Proceedings of International Web Science Conference*. ACM, New York, NY, 21.

[34] Jannis Kallinikos, Aleksi Aaltonen, and Attila Marton. 2013. The ambivalent ontology of digital artifacts. *MIS Quarterly* 37, 2 (2013), 357–370.

[35] James M. Kendra and Tricia Wachtendorf. 2003. Reconsidering convergence and converger legitimacy in response to the World Trade Center disaster. *Research in Social Problems and Public Policy* 11, 1 (2003), 97–122.

[36] Yoonsang Kim, Rachel Nordgren, and Sherry Emery. 2020. The story of Goldilocks and three Twitter's APIs: A pilot study on Twitter data sources and disclosure. *International Journal of Environmental Research and Public Health* 17, 3 (2020), 864.

[37] Olivia Peterkin. 2012. It's been a month since Hawaii's false missile alert. here's what's changed. https://www.hawaiinewsnow.com/story/37496908/recap-whats-happened-in-the-month-since-the-false-missile-alert/.

[38] Marina Kogan, Leysia Palen, and Kenneth M. Anderson. 2015. Think local, retweet global: Retweeting by the geographically-vulnerable during Hurricane Sandy. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'15)*. 981–993.

[39] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web (WWW'10)*. ACM, New York, NY, 591–600.

[40] Jessica Li and H. Raghav Rao. 2010. Twitter as a rapid response news service: An exploration in the context of the 2008 China earthquake. *Electronic Journal of Information Systems in Developing Countries* 42, 1 (2010), 1–22.

[41] Ingrid Lunden. 2020. Twitter prioritizes blue-check verifications to confirm experts on COVID-19 and the novel coronavirus. *TechCrunch*. Retrieved January 24, 2023 from https://techcrunch.com/2020/03/21/twitter-prioritizes-blue-check-verifications-to-confirm-experts-on-covid-19-and-the-novel-coronavirus/.

[42] Jim Maddock, Kate Starbird, Haneen J. Al-Hassani, Daniel E. Sandoval, Mania Orand, and Robert M. Mason. 2015. Characterizing online rumoring behavior using multi-dimensional signatures. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'15)*. ACM, New York, NY, 228–241.

[43] Andrew Marantz. 2019. *Antisocial: Online Extremists, Techno-Utopians, and the Hijacking of the American Conversation*. Viking.

[44] Robert K. Merton. 1968. The Matthew effect in science: The reward and communication systems of science are considered. *Science* 159, 3810 (1968), 56–63.

[45] Laura K. Nelson. 2020. Computational grounded theory: A methodological framework. *Sociological Methods & Research* 49, 1 (2020), 3–42.
[46] Venkata Kishore Neppalli, Murilo Cerqueira Medeiros, Cornelia Caragea, Doina Caragea, Andrea H. Tapia, and Shane E. Halse. 2016. Retweetability analysis and prediction during Hurricane Sandy. In *Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM'16).*
[47] Netaustin. 2019. Catching up with the teenager who sent first tweet debunking last year's missile alert. *KHON2.* Retrieved January 24, 2023 from https://www.khon2.com/news/catching-up-with-the-teenager-who-sent-first-tweet-debunking-last-years-missile-alert/amp/.
[48] Onook Oh, Manish Agrawal, and H. Raghav Rao. 2013. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly* 37, 2 (2013), 407–426.
[49] Onook Oh, Chanyong Eom, and H. Raghav Rao. 2012. Collective sense-making through the Twitter service during the 2011 Egypt revolution. In *Proceedings of the International Conference on Information Systems (ICIS'12).*
[50] Onook Oh, Kyounghee Hazel Kwon, and H. Raghav Rao. 2010. An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti earthquake 2010. In *Proceedings of the International Conference on Information Systems (ICIS'10).* 7332–7336.
[51] Wanda J. Orlikowski. 1992. The duality of technology: Rethinking the concept of technology in organizations. *Organization Science* 3, 3 (1992), 398–427.
[52] Wanda J. Orlikowski. 2007. Sociomaterial practices: Exploring technology at work. *Organization Studies* 28, 9 (2007), 1435–1448.
[53] Leysia Palen. 2008. Online social media in crisis events. *EDUCAUSE Quarterly* 31, 3 (2008), 76–78.
[54] Leysia Palen and Sophia B. Liu. 2007. Citizen communications in crisis: Anticipating a future of ICT-supported public participation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 727–736.
[55] Leysia Palen and Sarah Vieweg. 2008. The emergence of online widescale interaction in unexpected events: Assistance, alliance & retreat. In *Proceedings of the 2008 ACM Conference on Computer-Supported Cooperative Work (CSCW'08).* 117–126.
[56] Leysia Palen, Sarah Vieweg, Sophia B. Liu, and Amanda Lee Hughes. 2009. Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. *Social Science Computer Review* 27, 4 (2009), 467–480.
[57] Sergio Picazo-Vela, Marilu Fernandez-Haddad, and Luis F. Luna-Reyes. 2016. Opening the black box: Developing strategies to use social media in government. *Government Information Quarterly* 33, 4 (2016), 693–704.
[58] Soumajit Pramanik, Qinna Wang, Maximilien Danisch, Jean-Loup Guillaume, and Bivas Mitra. 2017. Modeling cascade formation in Twitter amidst mentions and retweets. *Social Network Analysis and Mining* 7, 1 (2017), 1–18.
[59] Daniel M. Romero, Wojciech Galuba, Sitaram Asur, and Bernardo A. Huberman. 2011. Influence and passivity in social media. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases.* 18–33.
[60] Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts. 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311, 5762 (2006), 854–856.
[61] Tamotsu Shibutani. 1966. *Improvised News: A Sociological Study of Rumor.* Ardent Media.
[62] Irina Shklovski, Leysia Palen, and Jeannette Sutton. 2008. Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM Conference on Computer-Supported Cooperative Work (CSCW'08).* 127–136.
[63] Pamela J. Shoemaker. 1991. *Gatekeeping.* SAGE Publications.
[64] Pamela J. Shoemaker, Tim P. Vos, and Stephen D. Reese. 2009. Journalists as gatekeepers. In *The Handbook of Journalism Studies.* Routledge, 93–107.
[65] Herbert A. Simon. 1996. Designing organizations for an information-rich world. *International Library of Critical Writings in Economics* 70 (1996), 187–202.
[66] Alan T. Sorensen. 2007. Bestseller lists and product variety. *Journal of Industrial Economics* 55, 4 (2007), 715–738.
[67] Emma S. Spiro. 2020. Online communication by emergency responders during crisis events *The Oxford Handbook of Networked Communication.* Oxford University Press, 10.1093/oxfordhb/9780190460518.013.11
[68] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.
[69] Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M. Mason. 2014. Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston Marathon bombing. In *iConference 2014 Proceedings.*
[70] Kate Starbird, Grace Muzny, and Leysia Palen. 2012. Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground Twitterers during mass disruptions. In *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management (ISCRAM'12).* 1–10.

[71] Kate Starbird and L. Palen. 2010. Pass it on?: Retweeting in mass emergency. In *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management (ISCRAM'10)*.
[72] Kate Starbird and Leysia Palen. 2011. "Voluntweeters" self-organizing by digital volunteers in times of crisis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1071–1080.
[73] Kate Starbird and Leysia Palen. 2012. (How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising. In *Proceedings of the 2012 ACM Conference on Computer-Supported Cooperative Work (CSCW'11)*. 7–16.
[74] Kate Starbird, Emma Spiro, Isabelle Edwards, Kaitlyn Zhou, Jim Maddock, and Sindhuja Narasimhan. 2016. Could this be true? I think so! Expressed uncertainty in online rumoring. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 360–371.
[75] Bruce Sterling. 1992. Free as air, free as water, free as knowledge. In*Thinking Robots, an Aware Internet, and Cyberpunk Librarians: The 1992 LITA President's Program: Presentations by Hans Moravec, Bruce Sterling, and David Brin.* BMLA, 25–32.
[76] Bongwon Suh, Lichan Hong, Peter Pirolli, and Ed H. Chi. 2010. Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In *Proceedings of the 2010 IEEE 2nd International Conference on Social Computing*. IEEE, Los Alamitos, CA, 177–184.
[77] Twitter Support. 2020. PSA about what we're doing to Verify Twitter accounts that are providing credible updates around #COVID19: we are working with global public health authorities to identify experts and have already Verified hundreds of accounts, but there's more to do and we could use your help. *Twitter*. Retrieved January 24, 2023 from https://twitter.com/TwitterSupport/status/1241155701822476288.
[78] Jeannette Sutton, C. Ben Gibson, Nolan Edward Phillips, Emma S. Spiro, Cedar League, Britta Johnson, Sean M. Fitzhugh, and Carter T. Butts. 2015. A cross-hazard analysis of terse message retransmission on Twitter. *Proceedings of the National Academy of Sciences* 112, 48 (2015), 14793–14798.
[79] Chenhao Tan, Lillian Lee, and Bo Pang. 2014. The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 175–185. https://doi.org/10.3115/v1/P14-1017
[80] Rebekah Tromble, Andreas Storz, and Daniela Stockmann. 2017. We Don't Know What We Don't Know: When and How the Use of Twitter's Public APIs Biases Scientific Inference. Retrieved January 24, 2023 from https://mediawell.ssrc.org/citation/we-dont-know-what-we-dont-know-when-and-how-the-use-of-twitters-public-apis-biases-scientific-inference/.
[81] Zhongyu Wei, Yulan He, Wei Gao, Binyang Li, Lanjun Zhou, and Kam-Fai Wong. 2013. Mainstream media behavior analysis on Twitter: A case study on UK general election. In *Proceedings of the ACM Conference on Hypertext and Social Media*. ACM, New York, NY, 174–178.
[82] Karl E. Weick. 1993. The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly* 38, 4 (1993), 628–652.
[83] Michael J. Welch, Uri Schonfeld, Dan He, and Junghoo Cho. 2011. Topical semantics of Twitter links. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, 327–336.
[84] Joanne I. White and Leysia Palen. 2015. Expertise in the Wired Wild West. In *Proceedings of the 18th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'15)*. ACM, New York, NY, 662–675.
[85] Tom Wilson, Kaitlyn Zhou, and Kate Starbird. 2018. Assembling strategic narratives: Information operations as collaborative work within an online community. In *Proceedings of the 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'18)*. 183.
[86] Tim Wu. 2017. *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. Vintage.
[87] Dan Zarrella. 2009. The Science of Retweets. Retrieved January 24, 2023 from https://cdn2.hubspot.net/hub/53/file-13207809-pdf/docs/science-of-retweets-201003.pdf.
[88] Li Zeng, Kate Starbird, and Emma S. Spiro. 2016. #Unconfirmed: Classifying rumor stance in crisis-related social media messages. In *Proceedings of the 10th International AAAI Conference on Web and Social Media*.
[89] Qingyuan Zhao, Murat A. Erdogdu, Hera Y. He, Anand Rajaraman, and Jure Leskovec. 2015. Seismic: A self-exciting point process model for predicting tweet popularity. In *Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, New York, NY, 1513–1522.
[90] Lu Zhou, Wenbo Wang, and Keke Chen. 2016. Tweet properly: Analyzing deleted tweets to understand and identify regrettable ones. In *Proceedings of the 25th International Conference on World Wide Web (WWW'16)*. 603–612.
[91] Arkaitz Zubiaga, Maria Liakata, Rob Procter, Kalina Bontcheva, and Peter Tolmie. 2015. Towards detecting rumours in social media. In *Proceedings of the Workshops at the 29th AAAI Conference on Artificial Intelligence*.

Received 18 February 2021; revised 15 October 2022; accepted 8 November 2022