2021 IEEE 9th International Conference on Healthcare Informatics (ICHI) | 978-1-6654-0132-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICHI52183.2021.00061

Exploring Suicidality on Social Media: Qualitative Analysis of Twitter

I'sis Perry
Department of Public Health Sciences
University of North Carolina Charlotte
Charlotte, NC, United States
0000-0002-2289-2682

Albert Park PhD
Department of Software and Information Systems
University of North Carolina Charlotte
Charlotte, NC, United States
0000-0002-2419-0131

***Abstract*— This study qualitatively examines the major themes of suicidality expressed on Twitter and the overarching relationships between categories. We collected suicide related tweets using 4 sets of key terms provided in 5 previous studies, followed by qualitative analysis of 8,000 randomly selected tweets. We found a total of 30 categories, which were grouped into three major themes: issues that are related to personal, interpersonal, and societal. Everyday issues, including emotions and relationships were most abundantly discussed through the lens of personal (36%) and interpersonal (14%) perspectives. Individuals also expressed suicidality when discussing diverse societal issues (50%), such as health and politics, and show the importance of the well-being of the general public to suicidality at the individual level.**

Keywords—suicide; mental health; social media; twitter; qualitative analysis; communication; suicide prevention; tweets; open coding; suicidal ideation

I. Introduction

According to the World Health Organization, across the globe, 800,000 people a year die by suicide [1]. Internationally, it is the second leading cause of death for 15-29-year-olds [1]. Moreover, the total number of suicides increased by 6.7% from 1990 to 2016, and some countries may be underreporting their numbers for suicide, making the matter much worse than it appears [5, 46].

In the United States, the Center for Disease Control and Prevention (CDC) found suicide was the tenth leading cause of death overall in 2017. Furthermore, suicides were more than 2 times more common than homicides that year [4]. Currently, suicide is the second leading cause of death for 10-34 year-olds [4]. Even more concerning than the number of suicides per year is the rate at which suicide is increasing. The total suicide rate in the US rose 31% from 2001 to 2017 [4]. The rate reflects an increase in suicides among both genders [14, 24]. In 2018, suicide rates increased in every state, except Nevada [7]. Studies have shown that the rates are higher for men, non-Hispanic whites, less educated people, and children associated with poverty [4], [8], [9]. Due to the increase in suicides, the number of individuals who are engaged in self-harming behavior or dealing with suicidal ideation is becoming more important to researchers [10]. The suicide rate is growing over time and it needs to be addressed to minimize its effect.

As of 2018, 88% of teens in the United States have access to computers and 95% have access to smartphones [11]. The numbers only drop to 75% and 93%, respectively, for those living in households that make less than 30,000 a year [11]. This means the access to the internet, and by extension social media, is almost universal, especially with the prevalence of public Wi-Fi [12]. Despite there being no clear conclusion to social media's effects, 89% of teens surveyed by the Pew Research Center stated they were online several times a day or almost constantly [13]. Additionally, older demographics used social media at a high rate as well [14]. These numbers show that the use of social media is deeply embedded in our culture.

Numerous attempts have been made to find the relationship between social media use and mental health [15]–[19], including suicidology research that has assessed determinants of suicidal ideation based on geography, cyberbullying, and sub-culture involvement [20]–[23]. This is partially possible due to suicidal users having been shown to use social media sites to express their own suicidal thoughts, even moments prior to acting on the stated thoughts [24]. Analysis can be done to determine the content most likely included within their posts. Several studies have used social media as a comparison tool against national suicide trends [21].

The extensive research on the topic has also revealed positive and negative influences on suicidal behavior and intent [15, 16, 19, 35]. Suicidal users, whether or not engaging in self-harm activities, are found to be negatively influenced by forums or sites dedicated to sharing tips on methods, concealment of harmful behavior, graphic imagery, and pro-suicide propaganda [15, 35]. It has also been asserted that users that purposefully search for self-harm related or risky content online are more likely to exhibit suicidal behaviors and be involved in communities of users who are also involved in self-harming behaviors [3, 4, 36–39, 41, 56]. Therefore, knowing the suicide related content available or search results relating to suicide on a site could be valuable. Additionally, exposure to cyberbullies is linked to increasing suicidal behaviors [26]. Exposure to celebrity deaths also seems to influence negative behavior. One study found a higher frequency in posting and more negativity within posts in the time periods immediately following celebrity deaths [37]. Conversely, some suicidal users are positively influenced by finding a sympathetic community and resources for seeking help [15, 35]. Another study found that users who are experiencing greater levels of ideation and are at an increased risk for self-harm are more likely to seek out negative, suicide focused materials; whereas, lower-risk users are less likely to specifically search for suicide related content. Furthermore, higher risk individuals avoid beneficial content related to improving their mental health [27]. Finally, social media

978-1-6654-0132-6/21/$31.00 ©2021 IEEE
DOI 10.1109/ICHI52183.2021.00061

supports interaction with others, which may include suicidal individuals. Users become less suicidal when interacting with others that share an understanding of their feelings and encourage them to seek help; this is particularly true for users facing isolation in the real world [28]. It is still not completely understood whether engaging in online communication regarding mental health will lead to deeper ideation for users or sparking ideation in others [12, 54]. Using the knowledge of how suicidal users interact with the internet, researchers are hoping to adopt strategies to combat their feelings and provide them with the necessary assistance.

Many studies have used Twitter as a specific platform for examining suicidality on social media [21], [40]–[51]. The American Psychological Association defines suicidality as "the risk of suicide, usually indicated by suicidal ideation or intent, especially as evident in the presence of a well-elaborated suicidal plan" [52] Suicidal users are more likely than the general population to post suicidal content [40]. This allows investigators to use language to identify risk factors and group tweets suspected of being indicative of suicidal ideation [47]. Through the use of machine learning and regression analyses, experts can differentiate between suicidal and non-suicidal tweets and users [41], [51]. In addition, researchers can categorize the tweets by concern levels, which could help lead to further investigation of the users [49]. Moreover, research suggests that users experiencing suicidal ideation are likely to be involved in similar communities, proliferating tweets related to suicidal content [45]. This means that suicidal users are more likely than non-suicidal users to have other suicidal users following them and re-tweeting their messages [10, 43]. Furthermore, they are more likely to post content if they identify with a mental health group at a higher risk for suicidal behavior. This was observed in a study where researchers tracked the frequency of suicide related posts of self-identified Schizophrenic Twitter users. This research lays a foundation for identifying particular groups and people who might be more likely to suffer from suicidal ideation.

In order to assemble the pertinent datasets, researchers typically employ one or more of the following methods. First, they select terms that are the most relevant to the area of interest. They may also define filter terms or stop words in order to ensure the proper retrieval of appropriate data [21], [42]. Second, after using Twitter's application programming interface to collect tweets, coders conduct a manual review to confirm appropriateness of the dataset and to begin categorization [49], [51]. Finally, some researchers pre-define terms in order to create training datasets, which will be utilized within various computational techniques [45], [50]. There are limitations to each of these methods. Inadequate keyword selection may cause datasets to be filled with a higher than intended percentage of irrelevant content. Additionally, manual classification is time consuming and not scalable. Finally, the use of pre-defined labels may not allow the datasets to be holistically representative of the subject; missing labels may be unique to the social media platform or may have an influential but rare association to the subject matter.

Knowing that there are likely Twitter users experiencing suicidal ideation, it could be beneficial to attempt to identify the most at-risk population. Past literature shows that exposure to suicidal content and interaction with suicidal users can increase the likelihood of succumbing to suicidal behavior [5]. It has also shown that there may be an increase in followers with suicidality for users that exhibit their own ideation within their posts [50]. In order to have the most accurate results, social media investigators performing research based on users' content must evaluate their datasets by assessing noise and relevance to their research questions. We chose to perform a qualitative analysis to better understand the content of our datasets. Utilizing Twitter, we will specifically examine the following research questions (RQ):

RQ1: What are qualitatively identified themes from tweets containing suicidality?

RQ2: What are the characteristics of tweets obtained using suicide related search terms?

We restricted our analysis to publicly available discussion content. The study was determined was granted exemption from review by the University of North Carolina-Charlotte's Institutional Review Board.

II. METHODS

A. Dataset

Twitter is a micro-blogging social media platform that broadcast short messages called tweets. We identified previously published works that included a set of terms/phrases and filter terms/phrases indicative of suicidal thoughts [21], [42], [45], [49], [51]. We employed the API Tweepy and used keywords taken directly from the previous studies to collect relevant datasets representing suicidality and filter out data as described in respective manuscripts. We independently executed the collection process for each set of terms, and they were concurrently collected for 4 weeks from October 4, 2019 to October 31, 2019. For each set of search terms, we randomly selected 1,000 original tweets generated from the first 2 weeks of data retrieval and another 1,000 from the second half of the 4 weeks. Therefore, for each of the four sets of terms, we separately analyzed 2000 tweets total. Overall, we analyzed 8,000 tweets. As we performed our analysis, we maintained a separation in the resulting tweets.

Table I illustrates the stylistic and terminology differences of selected keywords from previous works [21], [42], [45], [49], [51].

TABLE I. Selected Example Search Terms from previous Work

Paper	Selected Example Search Terms
[45]	"Just want to sleep forever" "Can't do this anymore" "Tired of being alone"
[21]	"bullying" "feelings of isolation" "impulsiveness"
[42]	"self-harm" "hang myself" "want to die"
[49], [51],	"my suicide note" "not worth living" "sleep forever" "better off without me" "suicide plan"

B. RQ1: What are qualitatively identified themes from tweets containing suicidality?

To better understand suicidality discussed on Twitter, we conducted a thematic analysis following an open coding process. We labeled each tweet as ‘Suicide Reference’ or ‘Noise’. Suicide Reference tweets contained any indication of suicidality or mention of suicide. We considered Noise tweets to be free of suicidal content. We used Uniform Resource Locator (URL) links within tweets to identify the tweet content and relevancy to suicide.

In addition, we further qualitatively labeled each tweet based on its complete content using open coding process [38], [49]. We did not pre-identify labels before retrieving our data. Instead, we began by reading each tweet and then determining its individual labelling. This required identifying as many topics (e.g., school, job, food, music, homelessness, health, hope etc.) as possible relating to the tweet. Though we created labels based on our own understanding of the tweet’s content, a literature review was necessary prior to labelling to appropriately identify some suicidal content, such as warning signs [54], [55]. All tweets were examined multiple times to confirm that each tweet was labeled appropriately and with all possible labels. Labels that referenced suicide in a nonliteral manner were included within the ‘Suicide Reference’ grouping (e.g., ‘Career, Economic, Political, Spiritual Suicide’). After completing the initial review of the 8,000 tweets, we removed any labels that occurred less than 10 times and could not be grouped with any remaining labels. Due to relevancy, we kept labels relating to suicide.

We created **categories** by grouping similar labels. If labels were frequent and other labels were able to be grouped into them, we converted the label name into a category name. Similarly to the labelling, we determined the category names based on the patterns and relationships between the labels. For example, the labels ‘Signs’ and ‘Triggers’ were grouped into the category of ‘Warnings’, because someone exhibiting signs of suicidal ideation and someone interacting with a known trigger for their suicidal ideation are both warning signs for possible risk. When presenting examples, we modified tweets to protect the privacy of users; the wording was changed, but the meaning remained the same.

After verifying the categories, we sorted them into **themes**. In order to do this, we created groupings that were the most intuitive to us based on evaluating the trends in conversation. We found that tweets could fit into conversations on different interaction levels. In essence, our categories show what users are talking about, and our themes show how users are talking about those things. Our process is outlined in Figure 1.

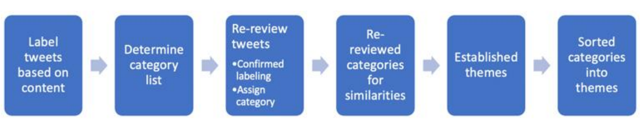

Figure 1. Category and Theme Creation

C. RQ2: What are the characteristics of tweets obtained using suicide related search terms?

We reviewed the individual categories and labels and assessed which were most frequent, within and between datasets. We began analysis by finding the ratio of Suicide Reference tweets within both the 1000-tweet and total 2000-tweet sets.

We identified labels that solely focused on an aspect of suicidality to grasp which topics most corresponded to suicide discussions within tweets. This is separate from when we grouped labels into categories; we wanted to isolate labels representing topics overtly engaging in discourse pertaining to suicide. Our purpose was to understand the underlying objective of suicide referencing tweets by examining their tone. In other words, we wanted to analyze the types of contributions users made towards the overall suicide discussion and whether those contributions tended to be more helpful or harmful to the discussion [49]. The labels were separated by whether their associated subject matters were conceptually viewed by the readers as positive, negative, or neutral regarding suicidality. Our concept of polarity was driven by wanting to know the overall intention behind the messages and their designated labels. We classified labels that were associated with tweets highlighting hope, suicide prevention, and awareness as positive. Negative labels referenced suicidal ideation, regret, completed suicides, suicide as acts of violence, methods, and calls for self-harm. Neutral labels were associated with tweets related to suicide facts, news, or general information, and nonliteral references to suicide. Additionally, neutral labels referenced content that could be interpreted as positive or negative depending on the context. We viewed negative labels as greater risk due to their overall expression of ideations or attempts.

III. Results

A. RQ1: What are qualitatively identified themes from tweets containing suicidality?

We designated 68% of tweets (n=5467) as ‘Noise’ and 31% of tweets (n=2466) as ‘Suicide Reference’. Of these tweets, 69 were ‘Indirect Suicide References’ and 4 were labeled as an ‘Incidental Suicide Reference’. Tweets with indirect suicide references did not concretely mention suicide but intimate the user deals with ideation. The tweet may have been as follows: “what if I never woke up again?” We defined an ‘Incidental Suicide Reference’ as a reference to the user taking their own life, not due to a desire to die, but rather an accident. For example, “I’d be so scared in a situation like that. I’d probably accidently kill myself.” The remaining 1% of tweets (n=67) were labeled possible, for tweets with potential ideation; these were unclear due to missing context. Table II summarizes the number of ‘Suicide Reference’ and ‘Noise’ related tweets, by the 1000 tweet datasets created from the search terms.

TABLE II. Dataset Summary

Dataset	Tweet Dates	Suicide Reference	Noise
Search Terms 1a	Oct 4-Oct 17, 2019	286	703
Search Terms 1b	Oct18-Oct 31, 2019	327	661
Search Terms 2a	Oct 4-Oct 17, 2019	64	933
Search Terms 2b	Oct18-Oct 31, 2019	64	933
Search Terms 3a	Oct 4-Oct 17, 2019	442	549
Search Terms 3b	Oct18-Oct 31, 2019	419	574
Search Terms 4a	Oct 4-Oct 17, 2019	444	540
Search Terms 4b	Oct18-Oct 31, 2019	420	574

We recorded a total of 179 labels. Table III provides examples of how the tweets were labeled, especially with consideration to having multiple labels per tweet. We did not implement a limit to the number per tweet; however, our results included a minimum of 1 label and a maximum of 15 labels.

TABLE III. Selected Labels (Modified) Examples

Tweet	Labels
On everything I'm not suicidal. I just think everything around me is falling apart and trying makes it worse. I'm absolutely trapped.	suicide reference-anti suicide-personal -trapped
Please read this. We need to pray for the Greens. We'll launch our safe house ministry soon. An abused 13-year-old dies by suicide	suicide reference-religious-conversation-suicide news-suicide awareness-specific-physical health-safety-age-abuse
I can sense the suicidal thoughts coming again...	suicide reference-personal-ideation
Democrats really try to tell religions how to practice their faith. I don't like it.	noise-political-religious

We sorted the labels into 30 categories. The number of tweets associated with each category ranged from 38 ('Life and Death') to 3359 ('Feelings'). After establishing the categories, we ensured each tweet was distinctly labeled as a 'Suicide Reference' or 'Noise'. We then created 3 themes: Personal, Interpersonal, and Societal by identifying commonalities in the focus of our categories.

The Personal theme dealt with the user's own feelings, thoughts, inclinations, etc. This theme also included categories created with labels typically seen as particular to an individual, such as 'Anniversary', which represents tweets that communicate the date of someone's completed suicide. It is important to note that we originally used Personal as a label for users who referenced suicidality pertaining to themselves. The Interpersonal theme encompassed categories that show interaction between multiple people. The Societal theme grouped together content related to society.

There were 8,179 Personal tweets, 3,274 Interpersonal tweets, and 11,376 Societal tweets (Table IV). These totals are higher than the reviewed tweet amount (i.e., 8,000) because the categories were not mutually exclusive and there were often multiple categories and labels per tweet. Only the groupings 'Suicide Reference' and 'Noise' were mutually exclusive, but they were not included in the theme consideration. Possible was excluded from the sort.

TABLE IV. Number of Tweets Per Major Groupings

Interpersonal	Personal	Societal
3274	8179	11376

Table V provides examples of some of the more prominent categories. Within the table, a count, theme, and definition of an example label for the chosen categories are included. Table III demonstrates how we labeled each tweet; Table V demonstrates how we defined categories and placed them into the broader themes. 'Suicide Focus' is Personal because associated tweets describe why users suffer from ideation. 'Health' is Societal because tweets reference topics that have an impact beyond the user and the user's relationships. Interpersonal includes categories that are contingent upon relationships between multiple people. For example, every tweet falling within 'Negative Relationships' details an adverse interaction between two or more people.

TABLE V: Examples of Categories

Label-Category (Sub-category)	Theme
Cause-Suicide Focus (Efforts) Referencing possible cause of suicide or ideation	Personal (298)
Death Reference-Health Referencing terms related to death	Societal (2160)
Bullying-Negative Relationships Referencing acts of bullying	Interpersonal (836)

Some categories (i.e., 'Suicide Focus', 'Feelings', 'Communication') diverge into sub-categories. The category 'Suicide Focus' (n=2628) incorporated labels that were defined by users' concentration on suicide within their messages. These messages expressed beliefs, risk, or history related to suicide. This category was further sorted into 4 sub-categories: 'Status' (n=1421), 'Ideology' (n=40), 'Against' (n=317), and 'Efforts' (n=850). 'Status' tweets discussed a user's current inclinations towards suicide or suicidal ideation. 'Ideology' tweets provided insight into what a user believes the reasons or purpose for suicide. 'Against' tweets pertained to a user's stance against suicide. 'Efforts' tweets contained information about a user's past or future attempts, including methods or regrets.

'Feelings' (n=3359), a category focused on an individual's emotional state and belief about themselves and their situation, was also sorted into sub-categories: 'Searching' (n=91) and 'Mental State' (n=3268). 'Searching' described a user's desire for escape from their current circumstances or feeling lost or trapped. 'Mental State' tweets included information about the user's emotions, lack thereof, or negative thoughts about themselves (e.g., worthlessness).

'Communication' (n=972) had two sub-categories: 'Figure of Speech' (n=939) and 'Style' (n=33). 'Communication' represented a grouping of labels related to unique diction and writing style. 'Figure of Speech' represented a compilation of labels that incorporate the word suicide without indicating literal suicidality (e.g., 'Career Suicide', 'Economic Suicide'). 'Style' represented a grouping of labels that reflected writing style of tweets (e.g., 'Questions', 'Quotes').

Figure 2 represents the final outcome of the sorting. We found that within Personal the related categories were: 'Risk Factors', 'Suicide Focus', 'Mental Health', 'Help', 'Rest', 'Feelings', and 'Life and Death'. Interpersonal had a focus on 'Negative Relationships', 'Positive Relationships', 'Warnings', 'Suicide Involvement', and 'Relations'. Societal tweets had the categories: 'Health', 'Entertainment', 'Mental Health Events', 'Politics', 'Work', 'Technology', and 'Other'.

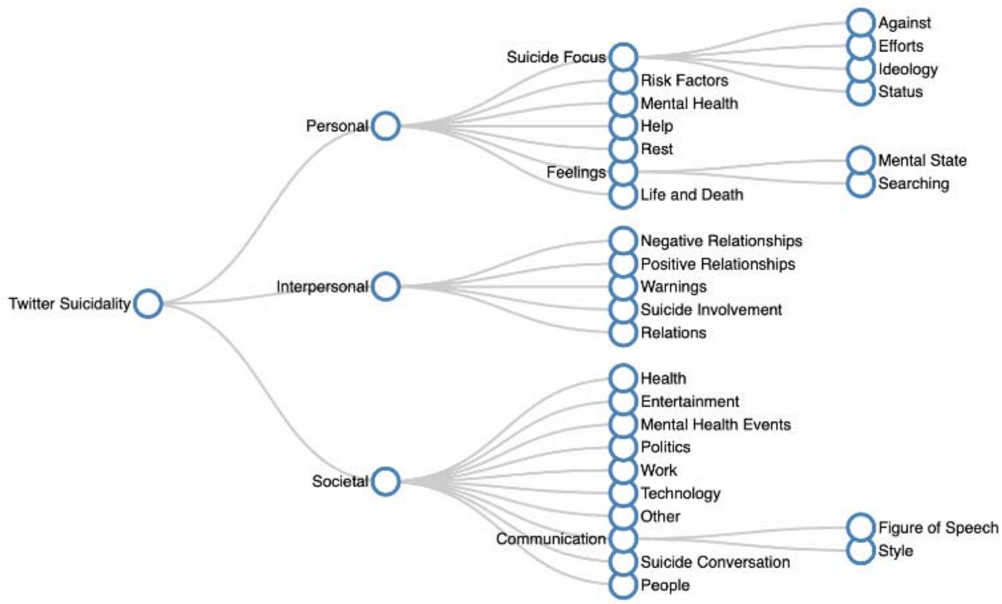

Figure 2: Sorting of Categories and Themes

B. RQ2: What are the characteristics of tweets obtained using suicide related search terms?

Table VI provides the top 5 labels per combined (2000 tweet) dataset. These are the labels that represent the overall narrative of individual datasets. Although the same labels were used between datasets, each set of search terms produced different amounts of each label. 'Political' and 'Relationships' were top five labels in each dataset. The datasets did not have even distribution of labels within or between datasets. In "Search Terms 1" the top label was associated with 346 tweets. The fifth highest label, 'Personal', was associated with 198 tweets. In "Search Terms 2" the top label was associated with 1037 tweets. The fifth highest label, 'Emptiness', was associated with 161 tweets.

TABLE VI. Top 5 Labels per Search Terms

Datasets (10/4/19-10/31/19)	Top 5 Labels
Search Terms 1	Political (346) Relationships (283) Death Reference (270) Self-Hate (198) Personal (198)
Search Terms 2	Emotions (1037) Sad (949) Relationships (431) Political (185) Emptiness (161)
Search Terms 3	Death Reference (930) Personal (376) Political (351) Relationships (376) Conversation (242)
Search Terms 4	Death Reference (865) Political (379) Personal 9270) Response (225) Relationships (224)

The top five categories were: 'Feelings' (n=3359), 'Health' (n=2884), 'Suicide Focus' (n=2628), 'Politics' (n=1831), 'Relations' (n=1576). The five with the lowest associations were: 'Rest' (n=308), 'Help' (n=116), 'Mental Health Events' (n=94), 'Warnings' (n=55), and 'Life and Death' (n=38).

Table VII shows our breakdown in polarity for labels relating to suicide. Only 60 labels met this criterion; we defined the labels based on their direct references to suicide. We found labels had suicide content that was positive, negative, or neutral (neither positive or negative or elements of both) in regard to the intention of the message. We found 13 positive labels which related to suicide decreasing efforts (e.g., "*operation shell shock ptsd awareness Military Suicide Prevention Program supporting the veterans and military crisis*" ['Suicide Prevention']). 'Suicide Prevention' and 'Anti-suicide' were the top two positive labels with 179 and 122 tweets containing the labels, respectively. Conversely, 29 had negative implications, including wanting or trying to commit suicide, as well as indicating a completed suicide (e.g., "*On my favorite holiday of the year I woke up unable to stop crying and wanting to die*" ['Ideation']). 351 tweets contained content relating to 'Ideation'. 137 tweets contained content relating to 'Methods' of suicide. The remaining 18 were considered neutral, the overwhelming majority of these were fall into this grouping because the context determined whether the content was meant positively or negatively (e.g., "*New numbers have been released on suicide rates*" ['Suicide News']).

TABLE VII. Non-overlapping Interpretation of Polarity of Suicide Themes from Twitter

Positive Labels (TOTAL: 13)	Anti-Joke (16) Anti-Suicide (122) Coping (30) Events (17) Help (5) Hotlines (29) Organizations (9)	Reasons (36) Reassurance (83) Remorse (1) Suicide Awareness (42) Suicide Prevention (179) World Suicide Prevention Day (2)
Neutral Labels (TOTAL: 18)	Anniversary (1) Career Suicide (8) Conversation (610) Economic Suicide (1) Future (12) Historical (45) Jokes (63) Nonserious (185) Personal (868)	Political Suicide (12) Questions (264) Response (607) Solution (1) Specific (247) Spiritual Suicide (1) Suicide News (159) Suicide Reference (2393) Target (58)
Negative Labels (TOTAL: 29)	Ambivalence (21) Assisted Suicide (14) Attempts (44) Cause (298) Commit Suicide (296) Fail (7) Farmers (16) Ideation (351) Imminent (20) Incidental Suicide Reference (4) Indirect Suicide Reference (69) Maliciousness (53) Mass Suicide (9)	Methods (137) Named (79) Purpose (3) Reactions (1) Sarcasm (99) Signs (2) Suicide Attacks (72) Suicide Cult (1) Suicide Mission (11) Suicide Notes (19) Suicide Pacts (2) Suicide Survivors (45) Suicide Threat (14) Suicide Watch (38) Ultimatum (9) Youth (50)

IV. DISCUSSION

In this paper, we qualitatively analyzed selected tweets from a collection gathered using suicide related search terms from previously published research papers [21], [42], [45], [49], [51]. We described suicidal content on Twitter and identified 3 major themes, which expanded to 30 categories. These findings are further discussed with respect to practical and public health implications.

A. Principle Findings

We found three major themes: Personal, Interpersonal, and Societal. The identified 30 categories were grouped into these themes, because issues were centered at those interaction levels. The themes were frequently discussed in the order of Societal, Personal, and Interpersonal. Although the users are creating the tweets, which is Personal by the nature of using social media, the outward interactions could also overtake the attention to self and the desire for suicide prevention and awareness. Therefore, tweets associated with Personal often had secondary labels that were associated with Societal (e.g., "I'm sad ['Emotions'; Personal] about their politics ['Political'; Societal] but I'm going to the movies ['Entertainment'; Societal] to forget about it"). Interpersonal possibly had the least number of tweets, because when relationships with others were described, the tweets generally included information about how the interactions affected them, leading to Personal labels. These patterns are partially reflected in literature regarding the linguistic style of suicidal individuals, which has been found to differ from non-suicidal individuals [56], [57]. Suicidal users are more likely to initiate messages on social media rather than replying to another user's posts, rely heavily on first person pronouns, and have global attributions [58].

Death Reference, Personal, Emotions, Relationships, and Political were among the most frequently expressed labels within the datasets. From our data, we found 5 potential explanations. First, the use of various forms of the word "death is pervasive in American English phrases (e.g., "dead wrong," "deadlock," "Work is killing me,"), which inflated the 'Death Reference' counts. Second, the datasets focus on suicide will make many tweets deeply personal (e.g., "*Depression had gotten me. I was having suicidal thoughts"*). Third, many tweets that were not about suicide still had references to mental health. Users' emotional states were often connected to their conversations on mental health (e.g., "*You can feel a breakdown coming. I zone out, get irritable, I can feel myself spiraling. And all I can do is watch and wait"*). Fourth, although the act of suicide is personal, there are many relational factors that increase or decrease risk of suicidality, including: ideation resulted from another person, interpersonal bonds that keep someone from committing suicide, and human desire to connect to another human. Finally, 'Political' may have persisted because the user population is represented by many Americans, and we live in a time when politics seem to permeate nearly every topic. Issues pertaining to farmers, military veterans, and LGBTQ citizens often focused on the particular populations elevated suicide risk and included language relating to legislation or political leanings (e.g., "*President X created millions more jobs then President Y threw a tantrum and farmers started committing suicide*"). We found many of the phrases that include the term suicide but are not meant to express literal suicidality (i.e., "*political suicide", "economic suicide*") within tweets that had a political tone (e.g., "*Senator X helped lead their political party to political suicide*"). Other phrases driving the interconnection between politics and suicide include metaphorical and literal attacks caused by politics (e.g., "*Senator Z went on a suicide mission. Everybody dies.*"). Politics was even included in tweets identified as noise (e.g., "*You tricked me into learning more about a politician. Interesting gun law reform proposal. My ex-marine husband and I go on dates to the shooting range.*")

Typically tweets focusing on suicide have negative connotations. However, our study also found positive social movement tweets related to suicide (22% of suicide specific labels), as shown in Table VII. A prime example is the conversation around Suicide Prevention and Reassurance. Previous literature computationally detecting suicidality using Twitter largely neglected this aspect of suicide discussion [47], [59]. Understanding that movement against suicide showcases that not every conversation regarding suicide must be problematic or considered a sign of suicidality. Additionally, we found a large number of neutral labels (30%), which included categories such as 'Solution', 'Personal' and 'Jokes'. However, tweets with a neutral label did not have to have a neutral tone. For example, a Historical tweet could read "I'm glad that suicide attempt failed" or "I wish my last attempt had worked."

Rarely occurring labels, for example, 'Rest', 'Help', 'Mental Health Events', 'Warnings', and 'Life and Death', discussed important experiences, such as getting treatment, identifying warning signs, and finding meaning in life (e.g., "*I can't take this anymore. I'm tired and no one understands me. I'm cutting* *myself…someone help…*"). Although these topics were not discussed as often as some of the other topics we identified, we believe they are central to the discussion of suicide [25], [60].

There were instances of slang within the tweets. Social media users range in age groups, meaning there are a variety of speech patterns and colloquial terms. "I'm dead" is a phrase commonly used by younger people to represent disbelief or finding humor in a situation. It does not represent literal death and usually is not used with an overly negative tone. This indicates that not every tweet captured our intended content. We included tweets that were collected because a search term served as slang in our 'Suicide Reference' grouping. However, when focusing on true suicidality, tweets composed using slang in a manner that lacks suicidality can be removed from analysis. When dealing with language and the culture surrounding social media sites, there will be unrelated posts, despite efforts to minimize noise. Our qualitative finding suggests that, on average, 68% of tweets from the four sets of search and filter terms were noise. One major reason for noise is that negative emotions or thinking, which are represented by some of the search terms, do not always equate to suicidality. We found many tweets that showed risk factors, such as depression, without any expression of suicidality (e.g., "*Dear PM my depression comes from your response to climate change and other issues. Your threat to young people is unacceptable.*"). We consider this noise because we were specifically looking for any indication of suicidality. Many tweets did not mention suicide or related content despite our selection of search terms based on words with a connection to suicide [21], [42], [45], [49], [51].

Tweets varied in their levels of associations with the search terms used, a partial explanation for the amount of noise. Numerous tweets did not have any easily identified search terms (e.g., "*Can't be tamed better than both*"). Some tweets indirectly referenced death and may have been retrieved because phrases like "*rest in peace*" often relate to death (e.g., "*I hope justice will be served and bodies returned to their families so they can finally rest in peace.*"). These did not refer to suicide. There were tweets that included the term death, but did not reference suicide (e.g., "*How dead do they need their enemies to be?").* In fact, it was possible for a tweet at this association level to express the urgency of life (e.g., "*Something that's not growing is dead. No? Life is urgent*"). Finally, there were instances of noise, referencing the user's death, but without the notion of real self-harm (e.g., "*Me playing dead in pools when I was little to see if anyone would care*"). Any reference to death caused by self-harm, even if it was not the user's, was marked as a Suicide Reference (e.g., "*Sayn just made a joke about Etika's Suicide"*). Therefore, search term sets that utilized terms that were less specific to suicide may have unintentionally extracted tweets that were important to mental health but did not definitively showcase suicidality, due to the unknown history of the individuals. Based on our analysis, we found a lesser number of users focused on suicide, and a greater number of users focused on other mental health issues, coping methods, emotional distress, and/or interpersonal interactions.

B. *Practical Implications and Future Directions*

Our findings suggest a number of methods to reduce noise from the dataset collected from social media posts. First, if possible, personal history should be accounted for in order to understand a user's relationship with suicidality. If profiles for suicidal users are identified and specifically selected for data retrieval, data may better reflect the general population of suicidal persons [61]. Second, we found a large variation of topics that are all related to suicide. This shows that users with different interests and backgrounds have a chance to interact with suicidal content. It also increases the chances for inclusion of unintended or unwanted topics. Therefore, retrieval should be as specific as possible. Third, we found that context matters when distinguishing slang or jokes from warning signs. Decoding another person's words can be confusing without context because the same phrase can be used literally or nonliterally (e.g., I want to de right now). Context should be accounted for when gathering data on specific topics. Contextual data cleaning and machine learning can be used to reduce noise and identify appropriate content [62]–[65].

Knowing the most prevalent content can propagate research by focusing on the interests already displayed by users. For example, based on our findings that death was referenced in over 25% of tweets, a focus could be on specifically finding messages that include the word "dead" in order to investigate which population is most likely to use it. A study could examine the context surrounding the usage and whether suicidal or non-suicidal users are more likely to casually include similar language within their posts. In a second example, if the conversation surrounding suicide seems to correlate highly with current coping strategies but hardly at all with seeking help, then researchers can focus on identifying what coping strategies are frequently used, which strategies are most effective, the population associated with specific strategies. Conversely, researchers can purposefully explore low correlating topics that they believe are important. Continuing the example, a researcher may ask why more users are not seeking help and could try to evaluate which users could be persuaded to seek treatment based on the characteristics of the users willing to include references to getting help. Consequently, the conversations around suicide may become more nuanced. Essentially, researchers are able to identify specific populations, needs of pre-selected populations, or trends relating to particular users. If more beneficial and appropriate content was available, higher risk users may not avoid them as easily, exposing them to potentially lifesaving information [27]. If health practitioners could promote topics more precisely on social media, they may have a better chance of positively influencing users to seek out help when experiencing ideation [15].

Categorization identified in this study can further enhance surveillance of suicidality on social media [66]. In order to further attempts at suicide prevention, experts can post prevention strategies in forums or within discussions pertaining to topics commonly referenced by suicidal users [67]. In this regard, public health workers can interact directly with users to answer questions, provide information, and provide prevention material [68]. Furthermore, realizing that at-risk users may have increased ideation after discussing mental health, practitioners may need to consider who should access their content [38]. If possible, in the future, access levels can be determined based on matching content with the risk profile of users. Twitter currently directly encourages users reported for showing signs of suicidality. Perhaps, prevention workers can collaborate directly with social media sites to automatically provide distinct prevention and coping strategies to reported users, based on their content and risk history, as an opt-in service.

Also, it has been found that average Twitter users are younger than the average United States citizen, although demographics for race, ethnicity, and gender were reflective of the population [69]. As previously discussed, suicide is the second leading cause of death for teenagers and young adults [4]. Therefore, selecting Twitter as a platform for our research potentially helped us comprehend content uniquely reflective of that specific at-risk population. Our findings suggest, social media researchers should investigate the demographics of users of the study's chosen platforms and compare these demographics to national trends so that they can know how closely their results apply to the real-world population of suicidal individuals or if the users they are studying are distinct. Public health practitioners can also use these findings by understanding that each social media site may have a unique user population needing tailored services or interactions. This will be needed if populations differ in the effectiveness of various prevention strategies. For example, knowing Twitter users are generally younger, public health workers can discuss coping strategies that are most useful for younger individuals at a higher rate than they discuss strategies useful for middle age individuals.

In addition, we found that not all discussion relating to suicide is intended to be negative, even if the subject matter is sensitive or the tone of a message appears to be negative. Suicide is a public health concern, and not only an individual's, as evidenced by the multitude of categories we identified, and the Interpersonal and Societal nature of many posts. Users were willing to discuss suicide in a range of topics. Public health workers may use this knowledge to post more freely and to encourage people to have open and honest conversations, without feeling burdened by previous levels of stigma.

We found it difficult to determine suicidality risk level performing a manual review. However, as aforementioned, machine learning can be used to determine individual risk. A theoretical model, such as *interpersonal-psychological theory of suicidal behavior* could be the basis for evaluating risk. User's profiles would be reviewed to assess perceived burdensomeness, their sense of belongingness, and their ability to die by suicide [70]. Perhaps, if profiles for suicidal users are automatically identified, researchers can shift their focus from reviewing and determining warning signs to finding which measures are most effective at prevention. In other words, there is an opportunity for a paradigm shift in suicide research; the perspective can change from concentrating on what increases the chances of negative outcomes to what increases the chances of positive outcomes. In turn, experts participating in general conversation regarding suicide may also utilize positive dialogue, potentially increasing engagement levels of suicidal individuals.

C. *Limitations*

There were multiple limitations within our research. First, the labels could be placed into multiple categories. For example, we sorted the Trauma label into the 'Mental Health' category. 'Trauma' may also be considered a risk factor [71]–[76]. Similarly, categories may fall into multiple themes. 'Depression Reference' may be 'Personal' if the user is speaking on their own feelings of depression; or 'Interpersonal' if the user is speaking about caring for a family member that has depression. However, we encounter such cases only a few times, thus we grouped them based on which theme the majority of the categories best aligned. Second, we were not always able to determine the appropriate labels for subjective tweets, due to the complexity of the English language, as well as the inability to entirely comprehend tone. However, researchers can use our results to understand which topics Twitter users most closely associated with suicide and understand possible intent of users. Further, language is constantly changing, and slang is prevalent on social media, hindering complete understanding of some tweets [77], [78]. In addition, social media sites, including Twitter, can foster a culture of trolling, a form of online harassment [79]. The harassers can go as far as threatening the lives of other users or encouraging others to commit acts of self-harm. The content of these tweets could potentially overlap with our search terms. Although we discuss this in terms of Twitter, these limitations apply to social media platforms in general. Third, our noise rate could be inflated, due to the fact we did not apply manual categorization [51] and machine learning classification [45]. Last, there were several nationally and globally recognized mental health and/or suicide events (e.g., National Mental Health Day (10/10/19), World Suicide Prevention Day (9/10/19)) occurring during the time period of our data collection. This could have skewed our numbers towards Suicide References and Societal tweets.

Suicide is a difficult subject to research. It is often a sensitive topic, the target population is vulnerable, and researchers cannot obtain firsthand knowledge from those who complete suicide [80]–[83]. Lacking a direct access to users limits our ability to know the clinical extent of their suicidality, though we manually examined each tweet. Future research can be conducted to compare clinically determined suicidal expression with expressions of suicidality on social media. In addition, some social media users who hesitate to seek help may not be forthcoming with their expression of their own suicidal ideations or tendencies, though their message may be influenced by their distress.

V. CONCLUSION

Our study is the first study to further qualitatively analyze datasets collected based on previously published articles. We found a total of 179 labels, which were grouped into 30 categories, which were further grouped into three major themes: Personal, Interpersonal, and Societal. Tweets had multiple categories and labels. Immediate risk levels for users were not easily identified when reviewed manually; however, we identified certain labels, that were more likely to be associated with a user's acknowledgement of personal ideation or general conversation regarding suicidality. Tweets addressing suicide and ideation did not always have a negative tone.

ACKNOWLEDGMENT

A.P. was supported, in part, by funds provided by the University of North Carolina at Charlotte.

REFERENCES

[1] World Health Organization, "Suicide data." [Online]. Available: https://www.who.int/mental_health/prevention/suicide/suicideprevent/en

[2] H. M. Orpana et al., "Global, regional, and national burden of suicide mortality 1990 to 2016: Systematic analysis for the Global Burden of Disease Study 2016," BMJ, vol. 364, 2019.

[3] J. M. Bertolote and A. Fleischmann, "Suicide and psychiatric diagnosis: a worldwide perspective.," World Psychiatry, vol. 1, no. 3, pp. 181–5, Oct. 2002.

[4] NIMH, "No Title." [Online]. Available: https://www.nimh.nih.gov/health/statistics/suicide.shtml#part_154968.

[5] S. C. Curtin, M. Warner, and H. Hedegaard, "Increase in suicide in the United States, 1999-2014. NCHS data brief, No. 241.," Natl. Cent. Heal. Stat., no. 241, pp. 1–8, 2016.

[6] H. Hedegaard, S. C. Curtin, and M. Warner, "Suicide Mortality in the United States, 1999-2017," 2018.

[7] CDC, "No Title," 2018. [Online]. Available: https://www.cdc.gov/vitalsigns/suicide/index.html.

[8] J. A. Hoffmann, C. A. Farrell, M. C. Monuteaux, E. W. Fleegler, and L. K. Lee, "Association of Pediatric Suicide With County-Level Poverty in the United States, 2007-2016," JAMA Pediatr., vol. 174, no. 3, p. 287, Mar. 2020.

[9] B. Han, P. S. Kott, A. Hughes, R. McKeon, C. Blanco, and W. M. Compton, "Estimating the rates of deaths by suicide among adults who attempt suicide in the United States.," J. Psychiatr. Res., vol. 77, pp. 125–33, Jun. 2016.

[10] J. Kidger, J. Heron, G. Lewis, J. Evans, and D. Gunnell, "Adolescent self-harm and suicidal thoughts in the ALSPAC cohort: A self-report survey in England," BMC Psychiatry, vol. 12, no. 1, p. 1, 2012.

[11] "Smartphone access nearly ubiquitous among teens, while having a home computer varies by income," 2018.

[12] Y. Yoshida, Takuya, Taniguchi, "Estimating the number of people using existing WiFi access point in indoor environment," 6th Eur. Conf. Comput. Sci. (ECCS '15), pp. 46–53, 2015.

[13] M. Anderson and J. Jiang, "Teens, Social Media & Technology 2018," 2018.

[14] A. Smith and M. Anderson, "Social Media Use in 2018," 2018.

[15] A. Park and M. Conway, "Longitudinal Changes in Psychological States in Online Health Community Members: Understanding the Long-Term Effects of Participating in an Online Depression Community.," J. Med. Internet Res., vol. 19, no. 3, p. e71, Mar. 2017.

[16] G. Gkotsis et al., "The language of mental health problems in social media," in Proceedings of the 3rd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2016, pp. 63–73.

[17] A. Park and M. Conway, "Harnessing Reddit to Understand the Written-Communication Challenges Experienced by Individuals With Mental Health Disorders: Analysis of Texts From Mental Health Communities.," J. Med. Internet Res., vol. 20, no. 4, p. e121, Apr. 2018.

[18] A. Park, M. Conway, and A. T. Chen, "Examining Thematic Similarity, Difference, and Membership in Three Online Mental Health Communities from Reddit: A Text Mining and Visualization Approach.," Comput. Human Behav., vol. 78, pp. 98–112, Jan. 2018.

[19] D. Mowery, A. Park, M. Conway, and C. Bryan, "Towards automatically classifying depressive symptoms from Twitter data for population health.," in Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media, 2016, pp. 182–191.

[20] K. J. Inder et al., "Determinants of suicidal ideation and suicide attempts: Parallel cross-sectional analyses examining geographical location," BMC Psychiatry, vol. 14, no. 1, 2014.

[21] J. Jashinsky et al., "Tracking suicide risk factors through Twitter in the US," Crisis, vol. 35, no. 1, pp. 51–59, 2014.

[22] M. A. Hughes, S. F. Knowles, K. Dhingra, H. L. Nicholson, and P. J. Taylor, "This corrosion: A systematic review of the association between alternative subcultures and the risk of self-harm and suicide," Br. J. Clin. Psychol., vol. 57, no. 4, pp. 491–513, 2018.

[23] A. John et al., "Self-Harm, Suicidal Behaviours, and Cyberbullying in Children and Young People: Systematic Review.," J. Med. Internet Res., vol. 20, no. 4, p. e129, 2018.

[24] X. Huang, L. Xing, J. R. Brubaker, and M. J. Paul, "Exploring Timelines of Confirmed Suicide Incidents Through Social Media," Proc. - 2017 IEEE Int. Conf. Healthc. Informatics, ICHI 2017, pp. 470–477, 2017.

[25] A. Marchant et al., "A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown," PLoS One, vol. 12, no. 8, p. e0181722, 2017.

[26] K. Daine, K. Hawton, V. Singaravelu, A. Stewart, S. Simkin, and P. Montgomery, "The power of the web: a systematic review of studies of the influence of the internet on self-harm and suicide in young people.," PLoS One, vol. 8, no. 10, pp. 1–6, 2013.

[27] C. Goldsmith, J. L. Donovan, D. Gunnell, L. Biddle, and J. Derges, "Using the internet for suicide-related purposes: Contrasting findings from young people in the community and self-harm patients admitted to hospital," PLoS One, vol. 13, no. 5, p. e0197712, 2018.

[28] T. Durkee, G. Hadlaczky, M. Westerlund, and V. Carli, "Internet pathways in suicidality: a review of the evidence.," Int. J. Environ. Res. Public Health, vol. 8, no. 10, pp. 3938–52, 2011.

[29] H. Sueki, "Association between deliberate self-harm-related Internet searches and the mental states and lifetime suicidal behaviors of Japanese young adults," Psychiatry Clin. Neurosci., vol. 66, no. 5, pp. 451–453, 2012.

[30] B. Mars et al., "Exposure to, and searching for, information about suicide and self-harm on the Internet: Prevalence and predictors in a population based cohort of young adults," J. Affect. Disord., vol. 185, pp. 239–245, 2015.

[31] A. M. Memon, S. G. Sharma, S. S. Mohite, and S. Jain, "The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of literature.," Indian J. Psychiatry, vol. 60, no. 4, pp. 384–392, 2018.

[32] K. Becker, M. Mayer, M. Nagenborg, M. El-Faddagh, and M. H. Schmidt, "Parasuicide online: Can suicide websites trigger suicidal behaviour in predisposed adolescents?," Nord. J. Psychiatry, vol. 58, no. 2, pp. 111–4, 2004.

[33] J. Bell, K. Mok, E. Gardiner, and J. Pirkis, "Suicide-Related Internet Use Among Suicidal Young People in the UK: Characteristics of Users, Effects of Use, and Barriers to Offline Help-Seeking.," Arch. Suicide Res., vol. 22, no. 2, pp. 263–277, 2018.

[34] K. Mok, A. F. Jorm, and J. Pirkis, "Who Goes Online for Suicide-Related Reasons?," Crisis, vol. 37, no. 2, pp. 112–20, Mar. 2016.

[35] T. Niederkrotenthaler, A. Haider, B. Till, K. Mok, and J. Pirkis, "Comparison of Suicidal People Who Use the Internet for Suicide-Related Reasons and Those Who Do Not.," Crisis, vol. 38, no. 2, pp. 131–135, Mar. 2017.

[36] K. Mok, A. F. Jorm, and J. Pirkis, "The perceived impact of suicide-related internet use: A survey of young Australians who have gone online for suicide-related reasons.," Digit. Heal., vol. 2, no. 0, p. 2055207616629862, 2016.

[37] M. Kumar, M. Dredze, G. Coppersmith, and M. De Choudhury, "Detecting Changes in Suicide Content Manifested in Social Media Following Celebrity Suicides.," HT ... Proc. ... ACM Conf. Hypertext Soc. Media. ACM Conf. Hypertext Soc. Media, vol. 2015, no. 24, pp. 85–94, Sep. 2015.

[38] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar, "Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media.," Proc. SIGCHI Conf. Hum. factors Comput. Syst. CHI Conf., vol. 2016, pp. 2098–2110, May 2016.

[39] L. Soldaini, T. Walsh, A. Cohan, J. Han, and N. Goharian, "Helping or Hurting? Predicting Changes in Users' Risk of Self-Harm Through Online Community Interactions," 2018.

[40] Y. Hswen, J. A. Naslund, J. S. Brownstein, and J. B. Hawkins, "Monitoring Online Discussions About Suicide Among Twitter Users

With Schizophrenia: Exploratory Study.," JMIR Ment. Heal., vol. 5, no. 4, p. e11483, Dec. 2018.

[41] S. R. Braithwaite, C. Giraud-Carrier, J. West, M. D. Barnes, and C. L. Hanson, "Validating Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality.," JMIR Ment. Heal., vol. 3, no. 2, p. e21, May 2016.

[42] J. Luo, J. Du, C. Tao, H. Xu, and Y. Zhang, "Exploring temporal suicidal behavior patterns on social media: Insight from Twitter analytics," Health Informatics J., 2019.

[43] B. J. F. Gunn and D. Lester, "Twitter postings and suicide: An analysis of the postings of a fatal suicide in the 24 hours prior to death," pp. 28–30, 2012.

[44] I. Cero and T. K. Witte, "Assortativity of suicide-related posting on social media.," Am. Psychol., Jun. 2019.

[45] G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield, "Analysing the connectivity and communication of suicidal users on twitter.," Comput. Commun., vol. 73, no. Pt B, pp. 291–300, 2016.

[46] A. Abboute, Y. Boudjeriou, G. Entringer, S. Bringay, and P. Poncelet, "Mining Twitter for Suicide Prevention Introduction et Motivations," Nat. Lang. Process. Inf. Syst. Vol. 8455 Lect. Notes Comput. Sci., pp. 250–253, 2014.

[47] S. J. Fodeh, J. Goulet, C. Brandt, and A.-T. Hamada, "Leveraging Twitter to better identify suicide risk," Proc. Mach. Learn. Res., vol. 69, no. 1, pp. 1–7, 2017.

[48] P. Burnap, G. Colombo, R. Amery, A. Hodorog, and J. Scourfield, "Multi-class machine classification of suicide-related communication on Twitter.," Online Soc. networks media, vol. 2, pp. 32–44, Aug. 2017.

[49] B. O'Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, "Detecting suicidality on twitter," Internet Interv., vol. 2, no. 2, pp. 183–188, 2015.

[50] P. Burnap, W. Colombo, and J. Scourfield, "Machine Classification and Analysis of Suicide-Related Communication on Twitter," in Proceedings of the 26th ACM Conference on Hypertext & Social Media - HT '15, 2015, vol. 2, pp. 75–84.

[51] B. O'Dea, M. E. Larsen, P. J. Batterham, A. L. Calear, and H. Christensen, "A Linguistic Analysis of Suicide-Related Twitter Posts.," Crisis, vol. 38, no. 5, pp. 319–329, Sep. 2017.

[52] "APA Dictionary of Psychology." [Online]. Available: https://dictionary.apa.org/suicidality.

[53] B. O'Dea, M. R. Achilles, M. E. Larsen, P. J. Batterham, A. L. Calear, and H. Christensen, "The rate of reply and nature of responses to suicide-related posts on Twitter.," Internet Interv., vol. 13, no. October 2017, pp. 105–107, Sep. 2018.

[54] A. F. for S. Prevention, "Risk Factors and Warning Signs." [Online]. Available: https://afsp.org/about-suicide/risk-factors-and-warning-signs/.

[55] S. Line, "Recognising suicide warning signs." [Online]. Available: https://www.suicideline.org.au/resource/recognising-suicide-warning-signs/.

[56] A. E. Aladağ, S. Muderrisoglu, N. B. Akbas, O. Zahmacioglu, and H. O. Bingol, "Detecting Suicidal Ideation on Forums: Proof-of-Concept Study.," J. Med. Internet Res., vol. 20, no. 6, p. e215, 2018.

[57] M. Chau, T. M. H. Li, P. W. C. Wong, J. J. Xu, P. S. F. Yip, and H. Chen, "Finding People with Emotional Distress in Online Social Media: A Design Combining Machine Learning and Rule-Based Classification," MIS Q., vol. 44, no. 2, pp. 933–955, Jun. 2020.

[58] A. Barak and O. Miron, "Writing Characteristics of Suicidal People on the Internet: A Psychological Investigation of Emerging Social Environments," Suicide Life-Threatening Behav., vol. 35, no. 5, pp. 507–524, Oct. 2005.

[59] G. Coppersmith, R. Leary, P. Crutchley, and A. Fine, "Natural Language Processing of Social Media as Screening for Suicide Risk," Biomed. Inform. Insights, vol. 10, p. 117822261879286, 2018.

[60] A. Costanza, M. Prelati, and M. Pompili, "The Meaning in Life in Suicidal Patients: The Presence and the Search for Constructs. A Systematic Review.," Medicina (Kaunas)., vol. 55, no. 8, Aug. 2019.

[61] A. Mbarek, S. Jamoussi, A. Charfi, and A. Ben Hamadou, "Suicidal Profiles Detection in Twitter," in Proceedings of the 15th International Conference on Web Information Systems and Technologies, 2019, pp. 289–296.

[62] M. Alipour-Langouri, Z. Zheng, F. Chiang, L. Golab, and J. Szlichta, "Contextual Data Cleaning," in 2018 IEEE 34th International Conference on Data Engineering Workshops (ICDEW), 2018, pp. 21–24.

[63] B. Calabrese, "Data cleaning," Encycl. Bioinforma. Comput. Biol. ABC Bioinforma., vol. 1–3, pp. 472–476, 2018.

[64] S. H. Yang, A. Kolcz, A. Schlaikjer, and P. Gupta, "Large-scale high-precision topic modeling on twitter," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 1907–1916, 2014.

[65] Q. Cheng and C. S. M. Lui, "Applying text mining methods to suicide research.," Suicide Life. Threat. Behav., vol. 51, no. 1, pp. 137–147, Feb. 2021.

[66] X. Dai, M. Bikdash, and B. Meyer, "From social media to public health surveillance: Word embedding based clustering method for twitter classification," in SoutheastCon 2017, 2017, pp. 1–7.

[67] J. Robinson et al., "Social media and suicide prevention: a systematic review.," Early Interv. Psychiatry, vol. 10, no. 2, pp. 103–21, 2016.

[68] J. Y. Breland, L. M. Quintiliani, K. L. Schneider, C. N. May, and S. Pagoto, "Social Media as a Tool to Increase the Impact of Public Health Research.," Am. J. Public Health, vol. 107, no. 12, pp. 1890–1891, 2017.

[69] S. Wojcik and A. Hughes, "Sizing Up Twitter Users," 2019.

[70] C. Chu et al., "The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research.," Psychol. Bull., vol. 143, no. 12, pp. 1313–1345, Dec. 2017.

[71] MCES, "Facts About Trauma & Suicide." [Online]. Available: http://www.mces.org/pages/suicide_fact_trauma.php.

[72] J. L. Gradus, M. K. Suvak, B. E. Wisco, B. P. Marx, and P. A. Resick, "TREATMENT OF POSTTRAUMATIC STRESS DISORDER REDUCES SUICIDAL IDEATION," Depress. Anxiety, p. n/a-n/a, Apr. 2013.

[73] D. M. LeBouthillier, K. A. McMillan, M. A. Thibodeau, and G. J. G. Asmundson, "Types and Number of Traumas Associated With Suicidal Ideation and Suicide Attempts in PTSD: Findings From a U.S. Nationally Representative Sample," J. Trauma. Stress, vol. 28, no. 3, pp. 183–190, Jun. 2015.

[74] H. Bermas and M. Sabet, "Correlation between Childhood Trauma and Suicidal Ideation by Mediation of Mental Pain and Object Relations یعوضوم طباور و یناور درد یرگیجنایم اب یشکدوخ یزادرپ هدیا و یکدوک یامورت یگتسبمه," vol. 9, no. 3, 2020.

[75] H. G. Prigerson et al., "Influence of traumatic grief on suicidal ideation among young adults," Am. J. Psychiatry, vol. 156, no. 12, pp. 1994–1995, 1999.

[76] N. H. Bahraini, G. K. Simpson, L. A. Brenner, A. S. Hoffberg, and A. L. Schneider, "Suicidal ideation and behaviours after traumatic brain injury: A systematic review," Brain Impair., vol. 14, no. 1, pp. 92–112, 2013.

[77] A. Mauranen and S. Vetchinnikova, Eds., Language Change: The impact of English as a lingua franca. Cambridge University Press, 2020.

[78] M. Zappavigna, Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web. .

[79] M. Sanfilippo, S. Yang, and P. Fichman, "Trolling here, there, and everywhere: Perceptions of trolling behaviors in context," J. Assoc. Inf. Sci. Technol., vol. 68, no. 10, pp. 2313–2327, Oct. 2017.

[80] L. Biddle et al., "Qualitative interviewing with vulnerable populations: Individuals' experiences of participating in suicide and self-harm based research," J. Affect. Disord., vol. 145, no. 3, pp. 356–362, Mar. 2013.

[81] S. Sharkey et al., "Ethical practice in internet research involving vulnerable people: lessons from a self-harm discussion forum study (SharpTalk)," J. Med. Ethics, vol. 37, no. 12, pp. 752–758, Dec. 2011.

[82] J. Bantjes and L. Swartz, "'What Can We Learn From First-Person Narratives?' The Case of Nonfatal Suicidal Behavior," Qual. Health Res., vol. 29, no. 10, pp. 1497–1507, Aug. 2019.

[83] C. A. Blades, W. G. K. Stritzke, A. C. Page, and J. D. Brown, "The benefits and risks of asking research participants about suicide: A meta-analysis of the impact of exposure to suicide-related content," Clin. Psychol. Rev., vol. 64, pp. 1–12, Aug. 2018.