



# A Critical Review of Text Mining Applications for Suicide Research

Jennifer M. Boggs<sup>1</sup> · Julie M. Kafka<sup>2</sup>

Accepted: 23 March 2022 / Published online: 26 July 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

**Purpose of Review** Applying text mining to suicide research holds a great deal of promise. In this manuscript, literature from 2019 to 2021 is critically reviewed for text mining projects that use electronic health records, social media data, and death records.

**Recent Findings** Text mining has helped identify risk factors for suicide in general and specific populations (e.g., older adults), has been combined with structured variables in EHRs to predict suicide risk, and has been used to track trends in social media suicidal discourse following population level events (e.g., COVID-19, celebrity suicides).

**Summary** Future research should utilize text mining along with data linkage methods to capture more complete information on risk factors and outcomes across data sources (e.g., combining death records and EHRs), evaluate effectiveness of NLP-based intervention programs that use suicide risk prediction, establish standards for reporting accuracy of text mining programs to enable comparison across studies, and incorporate implementation science to understand feasibility, acceptability, and technical considerations.

**Keywords** Suicide · Text mining · COVID-19

## Introduction

Suicide is a public health problem that impacts people around the globe [1]. The thoughts, emotions, and precipitating factors that contribute to suicide risk are complex and vary across cultures. Globally, prevention efforts have yielded mixed results, with some countries showing substantial reductions in rates from 2000 to 2015, while others (including the USA) saw dramatic increases in suicide during this time [2]. Suicide research has mostly examined structured (i.e., close-ended) data to understand risk factors (e.g., demographics, mental health diagnoses, substance use, social support) and to evaluate the impacts of prevention efforts (e.g., mental health treatment, restricting access to lethal means). While structured information

is highly valuable, it does not allow researchers to gain a deeper understanding of individual's lived experiences or to explore new risk factors that have not already been systematically recorded. There may be opportunities to identify new directions for prevention efforts by examining unstructured textual information.

There has been astronomical growth in the availability of electronic text in the past few decades, ranging from posts shared by individuals online through social media to clinical notes catalogued by providers in healthcare settings. Historically, qualitative methods using in-depth human review of small samples have been applied to provide rich and nuanced insights into behaviors, beliefs, or phenomena. Qualitative research, however, is not designed for processing large volumes of textual data, predicting outcomes longitudinally, or generating population-level inferences. Text mining provides the opportunity to use automated processes to systematically extract information from unstructured text [3]. Text mining can process thousands of records in seconds, rendering information as numeric variables which can then be used to predict or identify suicide risk. It incorporates many aspects of qualitative research, such as progressive, iterative steps to improve classification labels (e.g., history of suicidal ideation vs. no history), and in-depth human review of text using

---

This article is part of the Topical Collection on *Injury Epidemiology*

✉ Jennifer M. Boggs  
jennifer.m.boggs@kp.org

<sup>1</sup> Kaiser Permanente Colorado, Institute for Health Research, Aurora, CO, USA

<sup>2</sup> Department of Health Behavior, Gillings School of Global Public Health at University of North Carolina Chapel Hill, Chapel Hill, NC, USA

pre-defined rules that serve as the gold standard for accuracy checks.

For this review, we will include all types of text mining methods prevalently used in suicide research. Most studies use *Natural Language Processing*, commonly referred to as NLP, which encompass methods for [4] cleaning text, parsing terms, mapping grammatical and syntactical relations, and extracting information from qualitative text into coded discrete values, thus enabling quantitative data analysis.

Most text mining programs are either rule-based or data-driven. In either case, the end goal is usually to classify cases into groups (e.g., suicide risk vs. no suicide risk) based on some predetermined criteria. In rule-based programs, the researcher often identifies keywords or phrases that endorse a particular label, and then this logic is implemented deterministically. Term libraries or lexicons from public sources such as the United Medical Language System [5] or general sentiment (emotion) lexicons [6–8] are sometimes used, but often these need to be heavily edited for the specific context of suicide research [9]. Data-driven approaches use machine learning instead to create classification rules based on observed statistical associations between text-derived variables and other case information. Supervised machine learning entails having the researcher provide a priori labels for a subset of cases, and then the computer decides which text features distinguish each label. Unsupervised machine learning, which is less common in suicide research, is purely inductive and entails having the computer propose clusters or categories based on features in the data without a priori labels. This review focuses primarily on rule-based systems and supervised machine learning models.

Epidemiologic methods for diagnostic testing and qualitative research are commonly used to assess performance for both rule-based and data-driven approaches. Human review (i.e., human supplied labels) usually serves as the gold standard against which a text mining program is evaluated. Therefore, the program is only as accurate as the expertise and bias of the original human coders [10]. Evaluation of any text mining program should consider the background of the people conducting the labeling and accuracy checks. Similar to qualitative coding, rigorous methods should include double review, inter-rater reliability calculations, and a structured process for reconciling differences in labeling decisions between human reviewers [11]. Traditionally, performance measures are calculated based on the number of true positives (TP), false positives (FP), and false negatives (FN) when comparing the model's classifications against the human gold standard. Precision is analogous to positive predictive value ( $TP / TP + FP$ ), recall is also referred to as sensitivity ( $TP / TP + FN$ ), and F1 is the harmonic mean of precision and recall [12, 13]. None of these measures take true negatives (TN) into account.

The number of text mining programmatic tools has skyrocketed in the last decade including both open-source and proprietary products. Most of the scientific literature relies on open-source products. Python (an open-source programming language) has at least 10 different NLP libraries and tools to parse terms (e.g., part of speech tagging) [14]. Apache c-Takes [15, 16] is also a popular option. Python and Apache can be learned by students and data programmers, but they take time to master. These are not “out of the box” products that can produce seamless text extraction systems on day one. Proprietary products try to fill that niche [17] for industry customers (e.g., health care systems), but at the expense of transparency, which makes them unpopular with researchers.

Most suicide research that uses text mining examines online content, typically social media, comprising 45% of research articles from 2001 to 2019 [3], followed by electronic health records at 26% of articles. Another emerging area for suicide research is public health death records, such as the US Center for Disease Control's (CDC) National Violent Death Reporting System. For this review, we focus on three sources of data: social media/online content, electronic health records, and death records. For each area, we critically review research published from 2019 to 2021 and discuss both the practical utility of these findings for suicide prevention as well as directions for future research. Suicide notes and similar types of texts made up 19% of articles from 2001 to 2019 [3], but few articles emerged that focused on that content during this review for 2019–2021, so they were not included. Databases searched included the following: PubMed, Google Scholar, The Cochrane Library, Medline, PsychoINFO, PsychARTICLES, and ScienceDirect. Three categories of search terms were combined for each search: suicide terms (suicide, self-harm, mental health, depression, text mining terms); text mining terms (natural language processing, NLP, text mining); and content type terms (electronic health records, EHR, electronic medical records, EMR, treatment notes, psychotherapy notes, social media, Facebook, Twitter, Snapchat, death records, violent death reporting system).

## Electronic Health Records (EHR)

Electronic health records (EHR) contain structured, closed-ended data fields (e.g., demographic information, prescribed medications, diagnosis codes), as well as free-text fields. Free-text fields are rich and often contain more detailed information that is not captured elsewhere [18]. Most analyses of EHR records, however, only include structured data and therefore ignore about 80% of EHR content from free-text fields [19]. We review recent articles that evaluate the utility of adding text-derived information to structured EHR variables. Suicide research that uses text mining of EHRs has

had three main purposes in recent years: improving characterization of patient risk histories, identifying past treatments received, and predicting risk of a future suicide attempt.

### **EHR: Improving Characterization of Patient Risk Histories**

Properly characterizing patient histories using text fields in EHR data can address critically missing information for epidemiologic research. Text mining to find new mention of suicide ideation or suicide attempts that are not recorded elsewhere has been successful using both rule-based and machine-learning approaches [20, 21]. An NHS study in Britain of adolescents showed that relying on administrative codes alone would miss 83% of suicide risk histories [19]. Any approach to identify suicidal information in EHRs needs to consider three issues that cause false positives: references to history of ideation or self-harm that does not reflect current risk on the visit date (e.g. “patient cut herself 5 years ago”), negation terms such as “no suicide ideation,” and standard templates that include terminology like “past suicide attempts: none.”

Usually, studies that apply text mining to characterize patient histories are concerned with research study inclusion criteria and reducing recall bias. Currently, there are no direct clinical applications for this work, but text-mined information could be provided to support clinical decisions and patient risk assessment. Many medical records come with a manual version of this concept: a “search box,” for providers, but manual searching is time-consuming. Creating a succinct, close-ended field that summarizes suicidal history based on textual information could help clinicians more easily and effectively consider this during patient interactions.

### **EHR: Identifying Treatment Received**

Text mining of electronic health records can retrospectively identify the types of suicide prevention interventions that were delivered during routine patient care across larger samples. NLP programs have already been developed to identify when patients received lethal means counseling—an intervention when clinical providers counsel patients to voluntarily limit their access to firearms, medications, or other means for suicide. Our own work utilized text mining and illustrated a 75% reduced risk of suicide behavior for six months following receipt of lethal means counseling [22, 23]. There is potential to use this type of approach to evaluate other suicide prevention practices including safety or crisis response planning or other types of organization led initiatives that may have been recorded in text. Importantly, any observational study evaluating the impact of an intervention (determined by NLP or not) on suicide outcomes must carefully consider

selection bias and confounding by indication (treated patients are inherently higher risk than non-treated patients).

### **EHR: Predicting Suicide Risk**

Prediction models for suicidal behavior that use health records vary considerably in their accuracy [24–28]. Several studies have hypothesized that adding text-derived information to structured EHR data will increase predictive accuracy [29–31], with equivocal results. A case–control study of ~45,000 patients seen in emergency or inpatient settings for first-time intentional self-harm injury or poisoning at the University of Pittsburgh found a small increase of accuracy in risk prediction when text-mined data from health records were included [30]; however, it is unclear if performance gains were worth the programmatic complexity and additional resources required to include text-derived variables. A similar analysis of health records from UK’s NHS in ~18,000 patients found that free text about the past 30 days may be particularly informative for predicting medically treated self-harm behavior [31].

A significant drawback is that these prediction studies used hospital-based controls in their case–control designs. Case–control studies should choose controls that are representative (demographics, risk factors) of the sample that produced the cases. For suicide studies, hospital-based controls are convenient given the availability of EHR data, but hospital patients tend to have more medical morbidity than community controls [32]. Suicide risk is only moderately associated with medical comorbidity [33], so the characteristics of patients in the community are not well-represented with hospital-based controls. Retrospective cohort designs are a better choice when doing prediction work because participants are chosen uniformly and then followed forward to determine rates of the suicide behavior outcome. More work is needed to determine the “value-add” of creating text-derived variables for suicide risk prediction.

Clinical application of machine learning programs from EHRs that predict suicide risk (with and without text-mined information) are still in their infancy [34]. To our knowledge, with the exception of the Veterans Administration’s REACH VET [35] program (which does not include text-mined information), no other US-based health organization has published about successfully integrating suicide risk prediction models into routine clinical care, nor has there been extensive conversation about the ethics and appropriateness of these prediction systems [36]. One US-based study examined patient perspectives and found that patients believed using EHR information was acceptable for risk prediction purposes but feared a scenario where the computer model became the “holy grail” to assess suicide risk [37]. Patients wanted clinical (human) judgement to play a vital

role for risk determination and treatment planning. Adding text-derived variables to a prediction model requires significant additional computing effort, but, if this step improves accuracy and helps assuage concerns from patients (and providers) [38], then it may be worth it. Additional work is needed, however, to investigate the feasibility, costs, utility, and acceptability of text-informed suicide risk prediction modeling.

### **EHR: Summary**

Text mining EHR data holds promise for creating knowledge about suicide risk factors and evaluating the impact of prevention interventions like lethal means counseling on behavior outcomes, which can influence the field broadly. There is also considerable hype (and interest from some scientific organizations and funding agencies) [39] to further evaluate the utility of text-based suicide risk prediction models, but benefits are yet to be realized and results have been mixed.

The field also needs guidance from health system stakeholders on acceptable levels of performance, as there are no universally embraced standards [40, 41]. Furthermore, discussion is needed around evaluation practice given specific use cases. For example, if a healthcare system wanted to identify high-risk patients to implement an automated screening program for suicide ideation, they may want to maximize sensitivity over specificity; missing someone at risk for suicide is undesirable. On the other hand, documenting patient histories of suicide ideation for research purposes may benefit a more conservative approach that prioritizes specificity over sensitivity. Including clinical partners in this discussion is needed to explore the trade-offs, goals, and numerical benchmarks for successful model performance.

Most EHR studies focus on higher risk populations based on mental health indicators. This is a significant drawback considering that most people who die by suicide never seek mental health care in the year prior to their death [42]. It is vitally important that populations who are not engaged in mental healthcare are not neglected. Importantly, reporting on negative predictive value can also help assess the degree to which existing approaches miss high-risk patients. Medical risk factors (e.g., chronic conditions, opioid prescribing, traumatic brain injuries) [43, 44] have been associated with suicide in EHR studies, but few (if any) have included text-derived variables. Future text mining studies of EHRs should examine novel medical and social risk factors for suicide to predict risk in the general population. A new text mining tool called Moonstone was recently developed to help identify social risk factors including housing situation, living alone, and social support [45]. Considering novel risk factors is a key direction for future work.

## **Social Media Data**

Data from social media are abundant, updated frequently, and easy to access for research. Institutional Review Boards usually consider research on social media data exempt as long as data are public and individual users are not tracked or contacted [46], although ethical concerns have been raised about user's perceptions of privacy for the information they post online. Still, social media data are widely used in research, given that social media has an immense impact on communication globally, and can reflect common trends in our collective consciousness. We will focus on two main areas of text mining social media data for suicide research: individual risk detection and tracking population trends.

### **Social Media: Suicide Risk Detection (Individual Level)**

Social media data can provide timely information about suicide risk. There are already built-in alert systems on platforms like Facebook [47–49] to detect and respond to posts that contain suicide-related content. While this is promising, there is limited transparency about how these systems work and the types of errors these systems make. For example, evaluation metrics are not publicly available for Facebook's system that could be used to compare the proprietary approach to novel competitors [50]. A number of models have been proposed to detect suicide risk on Reddit and Twitter, although these have not been implemented system-wide [51, 52].

Researchers have begun exploring the utility of text-based suicide risk detection systems in other online settings too, including on blogs, online forums, and counseling environments. Some platforms use a "human in the loop" approach, where a computer model identifies individuals that have posted text that suggests potential suicidal ideation and then a human (often a trained counselor) is triaged in to address the situation. Other systems rely exclusively on automated response; for example, mental health chatbots sometimes manage self-harm talk by users without relying on humans for intervention [53].

In general, emerging research suggests that intervening online to address suicide risk can be accomplished in a way that is appropriate and acceptable to users. One qualitative study found that suicidal adolescents preferred having an automated system scan their online posts rather than have parents directly monitor their social media accounts; automated systems may infringe less on adolescents' sense of privacy, autonomy, and freedom of expression. Furthermore, the use of third-party monitoring systems entails no hands-on work or social media literacy by parents [52]. Still, more work is needed to determine what types of suicide

risk responses work. Honest conversations directly with suicidal users, their family, and/or guardians will be essential to guide this work.

While some suicide risk detection systems are already deployed, they have not yet been rigorously evaluated. Social media data lack verifiable outcome measures, and as a result, it remains unclear whether these systems can successfully link users to mental health services, influence underlying risk factors for suicide (e.g., loneliness, depression, social media addiction), or avert suicide attempts. Furthermore, very few text mining studies present background information about the users in their analytical samples (e.g., age, sex, race/ethnicity). This makes it challenging to understand what types of individuals share suicidal ideation or intent online and thus who can be reached using social media-based suicide risk detection systems.

Despite the absence of demographic or health history data in social media data, some researchers have extracted posts exclusively on certain sub-Reddits (e.g., for LGBTQ+ teens) [54] or have examined content that only contains focused keywords (e.g., epilepsy) [55] to study suicidality among people with particular identities or medical conditions. While this work is meaningful, it still does not paint a clear picture of who may be reached by online intervention and who will be missed. For example, one exploratory analysis of Reddit posts from suicide survivors found that the most frequent reported methods were drug overdose, hanging, and wrist cutting [56]. Conspicuously, firearms were not widely mentioned, although > 50% of suicides are firearm suicides [57]. Thus, social media provides an important, but incomplete picture of suicide risk.

### Social Media: Suicide Risk Trends (Population Level)

In addition to individual-level risk detection systems, social media data are used to monitor trends in population-level suicidality over time. In one study, researchers used NLP to examine how the Covid-19 pandemic impacted mental health. They analyzed posts from mental health support groups on Reddit and found the suicidality and loneliness clusters more than doubled in the number of posts during the pandemic compared to pre-pandemic [58]. Social media data has also been used to track how policy changes, viral posts, or celebrity self-harm events influence suicidality, anxiety, anger, and sadness [56, 59].

### Social Media: Summary

Overall, social media data are both richly detailed and widely available but have some important limitations. First, a large portion of the population does not regularly post or share content online. Second, social media data lacks verifiable

behavior and outcome information. Linking social media data to clinical or death data or enrolling social media users prospectively in experimental studies with their consent are potential next steps to evaluate models and determine whether they can have a verifiable impact on people's lives. Third, only a few studies have validated their text mining models using separate, representative datasets [52]. Even in these rare cases, descriptive information about users (e.g., gender, age, race/ethnicity) is often missing, which precludes examination of potential disparities in model effectiveness for individuals with different lived experiences. Evaluating models with consideration of demographic data is a critical next step. For greater transparency and to avoid exacerbating social biases [60], it will also be important to report whether models make different types of errors based on demographic and geographic characteristics of users, and how these errors could be corrected.

Methodologically, there has been sustained enthusiasm for research comparing the merits of different supervised machine learning algorithms to predict suicide risk using social media data. While this research has its place, it is already well-known that “more [training] data beats a cleverer algorithm [62].” Accordingly, the field should shift its focus away from exercises that strictly compare the performance of different learners (e.g., Random Forest versus XG Boost) and, instead, focus on the linguistically specific challenges of identifying and distinguishing suicide risk given different types of textual data. Some new models have been trained using textual data across different social media platforms [53], and some have even incorporated other text sources like suicide notes [50], which is a promising avenue for future work.

### Death Records

Due to limitations in research that exclusively uses suicide ideation or non-fatal self-harm outcomes to monitor suicide risk, researchers are beginning to use the US National Violent Death Reporting System (NVDRS [61]) to inform suicide prevention efforts. NVDRS uses information from coroners or medical examiner records (CME), law enforcement reports (LE), and death certificates to record suicide data. Trained NVDRS abstractors review these sources and summarize information using structured fields to capture decedent demographics, incident characteristics, and precipitating circumstances for the fatal event. They also summarize information about each death broadly using written text fields which are called “death narratives.”

Text mining can help extract information about novel risk factors for suicide from NVDRS death narratives. Researchers have already begun to examine suicides among older

adults (aged 55 years) and have found that  $n=305$  suicide cases (0.04%) in NVDRS were associated with driving cessation [61], while  $n=1037$  suicide deaths (2.2%) among older adults were associated with residential long-term care facilities [63]. Leveraging NVDRS death narratives thus can help elucidate new opportunities for screening and intervention, including providing more mental health support to older adults when they undergo significant life transitions. Other research to identify novel predictors for suicide using text mining with NVDRS death narratives, such as intimate partner violence or industry-specific job stress [64], is already underway [65, 66].

Recently, all 50 states were funded to collect data for NVDRS. For the first time, this will allow for detailed data on all suicide deaths in the USA. NVDRS compiles data only from a limited number of secondary sources and some indicators are underreported [62]. States often have access to certain key identifiers for suicide decedents. Thus, researchers could consider partnering with states to conduct deterministic or probabilistic linkages with other local data sources such as EHRs.

## Conclusions

Applying text mining for suicide research holds great promise, but additional work is needed to implement and evaluate proposed models, explore potential social biases in these models, and establish their acceptability for patients, providers, and online users.

In this review, we have identified variable levels of readiness to evaluate and/or implement text mining models in different settings. Interestingly, while social media studies are limited in rigor for predicting suicide, they have advanced further in terms of applying models to intervene on suicide risk across broad populations of users. More research is needed, however, to determine appropriateness and impact for these online risk detection systems. For EHR data, healthcare settings have not yet widely implemented risk detection models, and work is still underway to examine provider perspective, acceptance, and model utility [67]. Matching prediction models with evidence-based interventions for suicide prevention such as lethal means safety or engagement in psychotherapy is a logical next step.

Overall, more effort is needed to bridge the gap between text mining technical considerations and practical concerns. There are logistical considerations which must be addressed, such as computing requirements for running models within EHRs or on social media platforms. Recent social media research focuses primarily on methodological details without discussing practical implications. For example, we found no studies that explained how many false positives or false negatives might be identified on a daily basis when a social

media risk detection model was deployed. One of the obstacles to implementing text mining models on a larger scale is bridging the industry and research gap. Partnerships between researchers, social media companies, and/or healthcare organizations can help provide needed transparency and rigorous outcome evaluation. Furthermore, outside parties are needed to help social media companies specifically address aspects of their platform that increase self-harm risks to youth, especially given recent public scrutiny [68].

There is also a lack of commonly agreed upon norms for reporting results (e.g., STROBE, CONSORT) making comparison across text mining studies difficult. While performance measurement limited to precision and recall (sensitivity) may be sufficient in some fields, suicide researchers should also report negative predictive values and specificity. When we neglect these measures, we do not know who the prediction model overlooks.

Even when researchers use transparent methods and open-source programs, validation of programs across organizations or with new samples of data is lacking. Whether it is different social media platforms or different healthcare systems, each has a unique documentation culture that influences phrasing and word choice that text mining programs rely on. NLP toolkits do not always seamlessly transition across linguistic and social contexts, and extensive work is sometimes necessary to adapt these tools before they can be applied to new settings [69–71]. Thus, evaluating NLP tools using diverse, real-world datasets is a critical next step. Relatedly, model evaluations must present demographic stratifications to ensure the models do not perpetuate disparities [72], since data-driven approaches can replicate observed inequities [73, 74].

The last few years have seen significant advances in the application of text mining to categorize suicide risk, identify novel risk factors, catalogue patient suicide histories, document past interventions received, and track population-level trends in suicidality following major events (e.g., COVID-19). While additional work is needed to improve transparency, reporting norms, and translation for practice, text mining is poised to play a vital role in transforming how we study, assess, and respond to suicide risk.

## Declarations

**Conflict of Interest** The authors declare no competing interests.

## References

1. Naghavi M. Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the global burden of disease study 2016. *BMJ*. 2019;364.

2. Wu Y, Schwebel DC, Huang Y, Ning P, Cheng P, Hu G. Sex-specific and age-specific suicide mortality by method in 58 countries between 2000 and 2015. *Injury Prevention*. 2021;27(1):61–70.
3. Cheng Q, Lui CSM. Applying text mining methods to suicide research. *Suicide Life Threat Behav*. 2021;51(1):137–47.
4. Chowdhury GG. Natural language processing. *Annu Rev Inf Sci Technol*. 2003;37(1):51–89.
5. Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Res*. 2004;32(Database issue):D267–70. <https://doi.org/10.1093/nar/gkh061>.
6. Nielsen FÅ. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. *arXiv*. 2011; <http://arxiv.org/abs/1103.2903>. [accessed 2019-02-25].
7. Mohammad S, Turney P. Crowdsourcing a word-emotion association lexicon. *Comput Intell*. 2013;29(3):436–65.
8. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *J Lang Soc Psychol*. 2009;29(1):24–54.
9. Tadesse MM, Lin H, Xu B, Yang L. Detection of depression-related posts in reddit social media forum. *IEEE Access*. 2019;7:44883–93.
10. Liu T, Cheng Q, Homan CM, Silenzio V. Learning from various labeling strategies for suicide-related messages on social media: an experimental study. *arXiv preprint arXiv*. 2017:1701.08796.
11. McDonald N, Schoenebeck S, Forte A. Reliability and inter-rater reliability in qualitative research: norms and guidelines for CSCW and HCI practice. In: *Proceedings of the ACM on human-computer interaction*, 3(CSCW); 2019. p. 1–23.
12. Shunh KP. Accuracy, precision, recall or F1? Mar. 2018;15 <https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9>.
13. Dreisbach C, Koleck TA, Bourne PE, Bakken S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int J Med Inform*. 2019;125:37–46.
14. Top 10 Python NLP Libraries for 2019. *Analytics India Magazine*. May 24, 2019. <https://analyticsindiamag.com/top-10-python-nlp-libraries-for-2019/>.
15. Savova GK, Masanz JJ, Ogren PV, Zheng J, Sohn S, Kipper-Schuler KC, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): architecture, component evaluation and applications. *Journal of the American Medical Informatics*;16. Top Natural Language Processing Companies. [builtdin.com](http://builtdin.com), <https://builtdin.com/companies/type/natural-language-processing-companies>.
16. Top Natural Language Processing Companies. [builtdin.com](https://builtdin.com/companies/type/natural-language-processing-companies). <https://builtdin.com/companies/type/natural-language-processing-companies>
17. McKenzie K, Scott DA, Campbell MA, McClure RJ. The use of narrative text for injury surveillance research: a systematic review. *Accid Anal Prev*. 2010;42(2):354–63.
18. Kho AN, Rasmussen LV, Connolly JJ, et al. Practical challenges in integrating genomic data into the electronic health record. *BMJ*. 2013;15(10):772–8.
19. Carson NJ, Mullin B, Sanchez MJ, Lu F, Yang K, Menezes M, Cook BL. Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records. *PLoS One*. 2019;14(2):e0211116. <https://doi.org/10.1371/journal.pone.0211116>.
20. Fernandes AC, Velupillai S, Sanyal J, Stewart R, Chandran D. Identifying suicide ideation and suicidal attempts in a psychiatric clinical research database using natural language processing. *Sci Rep*. 2018;8(1):7426.
21. Boggs JM, Quintana LM, Powers JD, Beck A. Frequency of clinicians' assessments for access to lethal means in persons at risk for suicide. *Archives of Suicide Research*. 2020. <https://doi.org/10.1080/13811118.2020.1761917>.
22. Boggs, JM, Beck, A, Ritzwoller, DP, Anderson, H, Battaglia, C, Lindrooth, RC. A quasi-experimental analysis of lethal means assessment and risk for subsequent suicide attempts and deaths. *J Gen Intern Med*. 2020. <https://doi.org/10.1007/s11606-020-05641-4>.
23. Barak-Corren Y, Castro VM, Javitt S, et al. Predicting suicidal behavior from longitudinal electronic health records. *Am J Psychiatry*. 2017;174(2):154–62.
24. Simon GE, Johnson E, Lawrence JM, et al. Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. *Am J Psychiatry*. 2018;175(10):951–60.
25. Obeid JS, Dahne J, Christensen S, Howard S, Crawford T, Frey LJ, Stecker T, Bunnell BE. Identifying and predicting intentional self-harm in electronic health record clinical notes: deep learning approach. *JMIR Med Inform*. 2020;8(7):e17784. <https://doi.org/10.2196/17784>.
26. Cook BL, Progovac AM, Chen P, Mullin B, Hou S, Baca-Garcia E. Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. *Comput Math Methods Med*. 2016;8708434.
27. Pestian JP, Grupp-Phelan J, Bretonnel Cohen K, Meyers G, Richey LA, Matykiewicz P, et al. A controlled trial using natural language processing to examine the language of suicidal adolescents in the emergency department. *Suicide Life Threat Behav*. 2016;46(2):154–9.
28. Adamou M, Antoniou G, Greasidou E, Lagani V, Charonyktakis P, Tsamardinos I, Doyle M. Toward automatic risk assessment to support suicide prevention. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*. 2019;40(4):249–56. <https://doi.org/10.1027/0227-5910/a000561>.
29. Fuchiang R Tsui, Lingyun Shi, Victor Ruiz, Neal D Ryan, Candice Biernesser, Satish Iyengar, Colin G Walsh, David A Brent, Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts, *JAMIA Open*. 2021;4(1):00ab011,
30. Bittar A, Velupillai, S., Roberts, A., Dutta, R. International Medical Informatics Association and IOS Press. 2019. <https://doi.org/10.3233/SHT1190179>, <https://books.google.com/books?hl=en&lr=&id=m-9DwAAQBAJ&oi=fnd&pg=PA40&dq=Natural+language+processing+suicide+electronic+health+records&ots=qgrWsj5gtq&sig=LQ-NsYPTxlinVB7WHVjm0U2pyw4>.
31. Neupane B, Walter SD, Krueger P, Loeb M. Community controls were preferred to hospital controls in a case-control study where the cases are derived from the hospital. *J Clin Epidemiol*. 2010;63(8):926–31.
32. Ahmedani BK, Peterson EL, Hu Y, Rossom RC, Lynch F, Lu CY, et al. Major physical health conditions and risk of suicide. *Am J Prev Med*. 2017;53(3):308–15.
33. Levis M, Westgate CL, Gui J, Watts BV, Shiner B. Natural language processing of clinical mental health notes may add predictive value to existing suicide risk models. *Psychol Med*. 2021;51(8):1382–91.
34. Cannizzaro, K. REACH VET and the Possible Impact on Integrated Healthcare. <https://www.avapl.org/conference/pubs/2018%20Conference%20Presentations/Cannizzaro%20-%20slides%20-%20Open%20Plenary.pdf>.
35. McKernan LC, Clayton EW, Walsh CG. Protecting life while preserving liberty: ethical recommendations for suicide

- prevention with artificial intelligence. *Frontiers in Psychiatry*. 2018;9:650.
36. Yarborough BJH, Stumbo SP. Patient perspectives on acceptability of, and implementation preferences for, use of electronic health records and machine learning to identify suicide risk. *Gen Hosp Psychiatry*. 2021;70:31–7.
  37. Kessler RC, Bossarte RM, Luedtke A, Zaslavsky AM, Zubizarreta JR. Suicide prediction models: a critical review of recent research with recommendations for the way forward. *Mol Psychiatry*. 2020;25(1):168–79.
  38. Notice of Special Interest to Highlight Research Priorities for Risk Algorithms Applications in Healthcare Settings to Improve Suicide Prevention. NOT-MH-20-031. National Institute of Health. <https://grants.nih.gov/grants/guide/notice-files/NOT-MH-20-031.html>.
  39. Domingos P. A few useful things to know about machine learning. *Commun ACM*. 2012;55(10):78–87.
  40. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. *Clin Psychol Sci*. 2017;5(3):457–69.
  41. Ahmedani BK, Simon GE, Stewart C, Beck A, Waitzfelder BE, Rossom R, et al. Health care contacts in the year before suicide death. *J Gen Intern Med*. 2014;29(6):870–7.
  42. Boggs JM, Beck A, Hubley S, Peterson EL, Hu Y, Williams LK, et al. General medical, mental health, and demographic risk factors associated with suicide by firearm compared with other means. *Psychiatr Serv*. 2018;69(6):677–84.
  43. Simonetti JA, Piegari R, Maynard C, Brenner LA, Mori A, Post EP, et al. Characteristics and injury mechanisms of veteran primary care suicide decedents with and without diagnosed mental illness. *J Gen Intern Med*. 2020:1–7.
  44. Conway M, Keyhani S, Christensen L, South BR, Vali M, Walter LC, et al. Moonstone: a novel natural language processing system for inferring social risk from clinical narratives. *J Biomed Semant*. 2019;10(1):1–10.
  45. Moreno MA, Goniou N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. *Cyberpsychol Behav Soc Netw*. 2013;16(9):708–13.
  46. Suicide Prevention. Facebook. Meta. 2022; <https://www.facebook.com/safety/wellbeing/suicideprevention/>.
  47. Card, C. How Facebook AI helps suicide prevention. Sept. 10, 2018. <https://about.fb.com/news/2018/09/inside-feed-suicide-prevention-and-ai/>.
  48. Under the hood: Suicide prevention tools powered by AI. Feb. 21, 2018. <https://ai.facebook.com/blog/under-the-hood-suicide-prevention-tools-powered-by-ai/>.
  49. Zhang, Schoene, Ananiadou. Automatic identification of suicide notes with a transformer-based deep learning model. *Internet Interv*. 2021;25:100422.
  50. Skaik R, Inkpen D. Suicide ideation estimators within Canadian provinces using machine learning tools on social media text. *J Adv Inf Technol*. 2021;12(4).
  51. Roy A, Nikolitch K, McGinn R, Jinah S, Klement W, Kaminisky ZA. A machine learning approach predicts future risk to suicidal ideation from social media data. *NPJ digital medicine*. 2020;3(1):1–12.
  52. Liu X, Liu X, Sun J, Yu NX, Sun B, Li Q, Zhu T. Proactive suicide prevention online (PSPO): machine identification and crisis management for Chinese social media users with suicidal thoughts and behaviors. *J Med Internet Res*. 2019;21.
  53. Deshpande S, Warren J. Self-harm detection for mental health Chatbots. *Studies in health technology and informatics*. 2021;281:48–52. <https://doi.org/10.3233/SHTI210118>.
  54. Stevens HR, Acic I, Rhea S. Natural language processing insight into LGBTQ+ youth mental health during the COVID-19 pandemic: longitudinal content analysis of anxiety-provoking topics and trends in emotion in LGBTQ+ microcommunity Subreddit. *JMIR Public Health Surveill*. 2021;7(8):e29029. <https://doi.org/10.2196/2902>.
  55. Falcone T, Dagar A, Castilla-Puentes RC, Anand A, Brethenoux C, Valleta LG, Furey P, Timmons-Mitchell J, Pestana-Knight E. Digital conversations about suicide among teenagers and adults with epilepsy: a big-data, machine learning analysis. *Epilepsia*. 2020;61(5):951–8. <https://doi.org/10.1111/epi.16507>.
  56. Ambalavan AK, Moulahi B, Azé J, Bringay S. Unveiling online suicide behavior: what can we learn about mental health from suicide survivors of Reddit? *Studies in health technologies and informatics*. 2019;264:50–4. <https://doi.org/10.3233/SHTI190181>.
  57. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web-based Injury Statistics Query and Reporting System (WISQARS) [online]. (2005) [cited Year Month (abbreviated) Day]. Available from URL: [www.cdc.gov/injury/wisqars](http://www.cdc.gov/injury/wisqars).
  58. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19: observational study. *J Med Internet Res*. 2020;22(10):e22635. <https://doi.org/10.2196/22635>.
  59. Smith H, Cipolli W. The Instagram/Facebook ban on graphic self-harm imagery: a sentiment analysis and topic modeling approach. *Policy & Internet*; 2021.
  60. Turner NL, Resnick, P, Barton. G Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms The Brookings Institute May. 2019;22 <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms>.
  61. Centers for Disease Control and Prevention. National Violent Death Reporting System (NVDRS) Coding Manual Revised. [Online] 2021 National Center for Injury Prevention and Control, Centers for Disease Control and Prevention (producer). Available from URL: [www.cdc.gov/injury](http://www.cdc.gov/injury).
  62. Ko TM, Kalesnikava VA, Jurgens D, Mezuk B. A data science approach to estimating the frequency of driving cessation associated suicide in the US: evidence from the national violent death reporting system. *Front Public Health*. 2021;16(9):689967. <https://doi.org/10.3389/fpubh.2021.689967>.
  63. Mezuk B, Ko TM. Suicide among older adults living in or transitioning to residential long-term care, 2003 to 2015. *JAMA Netw Open*. 2019;2(6):e195627. <https://doi.org/10.1001/jamanetworkopen.2019.5627>.
  64. Davidson JE, Ye G, Parra MC, Chofflet A, Lee K, Barnes A, et al. Job-related problems prior to nurse suicide, 2003–2017: a mixed methods analysis using natural language processing and thematic analysis. *Journal of Nursing Regulation*. 2021;12(1):28–39.
  65. Engaging Machine Learning and Data Linkage to Understand Firearm Suicide Among Females. National Collaborative on Gun Violence Research: Our Grants. 2021. <https://www.ncgvr.org/grants/2021/engaging-machine-learning-and-data-linkage-to-understand-female-suicide.html>.
  66. Intimate Partner Violence and Firearm Suicide. Using text mining to bridge a critical gap in United States data systems. National Collaborative on Gun Violence Research: Our Grants; 2021. <https://www.ncgvr.org/grants/2021/intimate-partner-violence-and-firearmsuicide.html>.
  67. Yarborough B. H. Stakeholder perspectives on implementing suicide risk prediction models. <https://grantome.com/grant/NIH/U19-MH121738-01-7282>.
  68. Allyn, B. Here are 4 key points from the Facebook whistleblower's testimony on Capitol Hill. National Public Radio. Oct. 5,



2021. <https://www.npr.org/2021/10/05/1043377310/facebook-whistleblower-frances-haugen-congress>.
69. Sierra G, Andrade-Palos P, Bel-Enguix G, Osornio-Arteaga A, Cabrera-Mora A, García-Nieto L, Sierra-Aparicio T. Suicide risk factors: a language analysis approach in social media. *J Lang Soc Psychol*. 2021;0261927X211036171.
70. Singh J, Singh G, Singh R, Singh P. Morphological evaluation and sentiment analysis of Punjabi text using deep learning classification. *Journal of King Saud University-Computer and Information Sciences*. 2021;33(5):508–17.
71. Nikiforos S, Zervopoulos AD, Geramanis E, Toulakis A, Papamichail A, Triantafylloy D, et al. Suicidal tendencies prediction in Greek poetry. *Evol Syst*. 2021;12(1):57–68.
72. Coley RY, Johnson E, Simon GE, Cruz M, Shortreed SM. Racial/ethnic disparities in the performance of prediction models for death by suicide after mental health visits. *JAMA Psychiatry*. 2021.
73. Raji D. How our data encodes systematic racism. *MIT Technology Review*. <https://www.technologyreview.com/2020/12/10/1013617/racism-data-science-artificialintelligence-ai-opinion/>.
74. Williams BA, Brooks CF, Shmargad Y. How algorithms discriminate based on data they lack: challenges, solutions, and policy implications. *J Inf Policy*. 2018;8:78–115.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.