

From Durkheim to Machine Learning: Finding the Relevant Sociological Content in Depression and Suicide-Related Social Media Discourses



Júlia Koltai, Zoltán Kmetty, and Károly Bozsonyi

1 Introduction

The phenomenon of suicide is in the focus of social scientists since Durkheim (1897, 2002). In his classic work “Suicide” – he analyzed the spatial and temporal variation of suicide. He argued that in most cases, social causes could be found behind suicide variation. The topic was always in the center of scientific discourse, but after his work, even more attention was generated.

The work of Zoltán Kmetty was funded by the Premium Postdoctoral Grant of the Hungarian Academy of Sciences.

The work of Julia Koltai was funded by the Premium Postdoctoral Grant of the Hungarian Academy of Sciences.

J. Koltai (✉)

Centre for Social Sciences, Hungarian Academy of Sciences Centre of Excellence, Budapest, Hungary

Faculty of Social Sciences, Eötvös Loránd University, Budapest, Hungary

e-mail: koltai.julia@tk.mta.hu

Z. Kmetty

Centre for Social Sciences, Hungarian Academy of Sciences Centre of Excellence, Budapest, Hungary

Faculty of Social Sciences, Eötvös Loránd University, Budapest, Hungary

e-mail: kmetty.zoltan@tk.mta.hu

K. Bozsonyi

Faculty of Humanities and Social Sciences, Károli Gáspár University of the Reformed Church, Budapest, Hungary

e-mail: bozsonyi.karoly@kre.hu

© The Author(s) 2021

T. Rudas, G. Péli (eds.), *Pathways Between Social Science and Computational Social Science*, Computational Social Sciences, https://doi.org/10.1007/978-3-030-54936-7_11

237

In recent years, an increasing tendency of suicide has been observed in some western countries. In 2015, the female suicide rate in the UK reached its highest point of a decade.¹ Suicide rates also increased in the USA, especially in the young age cohorts. These tendencies drew the attention of social and medical scientists, and several studies were published, which attempted to provide an explanation for this phenomenon. One of the research teams who picked up on the increasing trends analyzed the online footprints of suicide and depression. Internet and social media sites provide new ways for people to express their positive feelings, but they are also platforms to express suicide ideation or depressed thoughts. Most of these contents are not notes about real suicides, but some of them are a cry for help. Suicide- and depression-related content varies among platforms, and it is not evident how a researcher can find these contents in mass data of social media. Accordingly, most of the studies, which focused on suicide, and used digital data for the research, used small sample sizes – big data approach is still uncommon in this field.

Digital data recording, social media, IoT, and digitalization, in general, produce such amount of data that was unimaginable in the past. Traditional surveys produced nicely structured and research question-related data, which was easy to analyze. These datasets are in sharp contrast with digital data, which are usually stored in highly unstructured formats and contain a lot of text data.

The noisiness, unstructuredness, and qualitative nature of this kind of data bring up new challenges for social researchers. It is far from evident how we can retrieve the relevant information from this vast amount of data. In our paper, we present possible solutions for this problem, applying a complex methodology of natural language processing, neural networks, and machine-learning algorithms.

We used the corpus of more than four million Instagram posts, related to mental health problems. However, this data is very diverse in content. It includes posts beyond those relevant from our theoretical point of view – trends of depression- and suicide-related content. Thus, after defining the initial corpus, we present two different strategies to find the relevant sociological content in the noisy environment of social media. The first approach starts with a topic modeling (Latent Dirichlet Allocation), resulting in an output that serves as the basis of a supervised classification method based on advanced machine-learning techniques. The other strategy is built on an artificial neural network-based word embedding language model. We primarily decided to use a word embedding model to validate the results of the topic model. Moreover, a word embedding approach can also help to discover the deeper understanding of main semantic patterns behind the posts we analyzed. The combination of the two methods provides a powerful framework for analyzing and interpreting the digital footprints of depression and suicide.

In the next part of the paper, we will present the most relevant findings of the studies, which deal with the digital representation of depression and self-harm. In the data and methods section, we will give an overview of the methods we used and

¹Suicides in the UK: 2015 registrations <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/suicidesintheunitedkingdom/2015registrations>

the data we collected for this study. In the empirical part, a very detailed explanation of the analysis process will be presented alongside the analysis of the results. In the final section, in addition to our conclusions, we will discuss the limitations of the paper and the possible application of the results in social media moderation practice.

2 Recent Studies of the Field

Recently several studies have been published investigating the possible relationships between the spread of suicides/mental problems in society and Internet use. Here we focus on those, which used social media data. In recent years, Twitter emerged as a valuable platform for examining the possible interrelationships of social media and suicide. A study conducted by Sueki (2015) examined the association between suicide-related tweets and the suicidal behavior of youngsters in Japan. The statistical results confirmed that tweeting “want to die” and “want to commit suicide” was significantly related to suicidal ideation and self-harming behavior. Although the findings are significant and relevant, the generalizability of the results is questionable, as the sample used in the study was not representative.

Braithwaite et al. (2016) created significant classifications for identifying participants at risk of suicide using a machine-learning method. However, their study can be criticized as it was based on a very low sample size. They used the data of only 17 at-risk participants for fitting their suicide-risk classification model. Kmetty et al. (2017) studied the temporal relationships between the moods of tweets and the number of committed suicides. They filtered more than 626 million tweets between February 1, 2012, and August 31, 2016, geolocated for the USA to identify those containing one of the following words, “depression,” “depressed,” “suicide,” “Prozac,” or “Zoloft.” The “bad mood index” created on the basis of the relative frequency of these words was found to be related to the distribution of actual suicides across the different days of the week. However, the monthly distribution of the index was unrelated to the seasonal pattern of suicides across the different months.

Besides Twitter, Instagram is also a very widely used social media platform. Brown et al. (2018) conducted their study on Instagram data. They analyzed German language Instagram posts starting from April 2016. After filtering out the hashtags, relevant for the purposes of the study, they selected 2826 pictures and related messages from 1154 users in total, portraying non-suicidal self-injuries (NSSI). They found that the number of such posts increased significantly in the evening. No differences were found between different days of the week, although on Sundays, there was a noticeable, but non-significant increase. They established that the number of messages had a low but significant positive correlation with the apparent severity of the injury pictured.

In June 2014, Moreno et al. (2016) collected 225 Instagram posts containing the “self-injury” hashtag. They found that because of the peculiar language habits of the users, the system used by Instagram for monitoring dangerous or harmful content

(“Content Advisory warning”) could not be trusted.² Reece and Danforth (2017) recruited 166 volunteers from Amazon’s Mechanical Turk crowd-working system, 71 of whom were diagnosed with depression. The researchers analyzed 13,184 Instagram photographs using both automatic image processing algorithms and human coders. In addition to the variables created on the basis of the characteristics of the pictures, the participants’ activities on Instagram, their popularity, and their social embeddedness were also taken into consideration. A statistical model was set up for classifying healthy/depressed and healthy/non-diagnosed depressed groups, which proved to be successful as they were able to classify users to the abovementioned groups based on their Instagram pictures and other related characteristics.

Andalibi et al. (2017) analyzed 788 Instagram posts and pictures, which included the hashtag #depression. They also examined the 1741 comments connected to these posts. The authors established a complex coding system to handle the emotional content of the posts, and they found that while posts express negative feelings and experiences, commenters often respond to those with positive feedback and support. Comments included more positive emotions than negative ones, whereas occurrences of negative emotion outnumbered positive ones in the posts.

Cheng et al. (2017) conducted an online survey among Weibo users ($n = 974$), where beyond a classic questionnaire participants’ Weibo posts published in the public domain were also downloaded. The associations between word counts of posts and suicide risk factors were examined and found that those with high suicide probability were marked by a greater total word count and higher usage of pronouns, prepend words, multifunction words, and lower usage of verbs. This latter finding is conflicting with the one of De Choudhury et al. (2016) as they found a contradictory relationship between the usage of verbs and suicidal ideation.

Although these studies used social media data to understand suicide, most of them were built on a small sample size. Thus, they could not provide an overview of how suicide and depression are represented overall in the online space. Our study takes one step forward and shows the main patterns of suicide- and depression-related contents of Instagram using a large amount of textual data.

3 Data and Methods

We used Instagram data collected by SentiOne, a web-based text analytics platform.³ The first post of the dataset was created on January 16, 2016, and the last one

²However, it is important to mention, that Instagram developed its system since 2014. See: <https://instagram-press.com/blog/2019/10/27/taking-more-steps-to-keep-the-people-who-use-instagram-safe/> (downloaded: November 8, 2019).

³SentiOne (www.sentione.hu) is an international social listening software, a content-based web analytics platform that covers and recognizes 30 languages all across Europe. It gathers, indexes, and analyses public online contents published anywhere on the web, starting with social media channels through blogs, forums, and websites. The interactive platform is built upon user-provided

on January 31, 2019. During this three-year period, Instagram data was collected every day with the same algorithm, through the application's API. The dataset contains all English language posts of this data collection.

From this database, we only used those posts, which include at least one hashtag and removed all other content, like simple texts and emojis from the posts. The approach is reasonable as post using hashtags become automatically searchable and can be accessed through the Instagram API. Posts without hashtags will only be searchable if the user set it public. It means that most of the Instagram posts, which are in the data warehouse of Sentione, are posts with hashtags, and there are only a few of that does not contain any. Hashtags are widely used on Instagram. People use hashtags to connect their pictures with different ideas or concepts; they are the simplified versions of associations and thoughts. One post can contain several hashtags and thus can provide a good overview of what words are linked in people's mind. Focusing only on hashtags simplifies the research as there is no need for text pre-processing, like stemming or lemmatization, or for deletion of articles, punctuation marks, and conjunctions as they do not appear in hashtags. Overall, using only hashtags does not only simplify the content but also makes the data more structured, and the linkages of words are still researchable.

In the first step, we selected those Instagram posts where specific hashtags appeared. We started with three keywords #depression, #depressed, and #suicide. Then, based on frequencies of the different words, which appeared in the selected posts, we extended the list of keywords with the following hashtags: #sucidal, #anxiety, #mentalhealth, #mentalillness, #selfharm, #sadness, #lonely, #broken, #cutting, #ptsd, #bipolar.

At this point, we did not use any stopwords to filter out the non-relevant content. Although stopwords could help us to clear out some of the non-relevant posts from the database, in this phase, it would have only served as a partial solution, because we could not define all the requisite stopwords in advance. That is why we decided to use a topic model approach at first to detect the relevant depression- and suicide-related contents.

The dataset, which was selected by the abovementioned hashtags, contained more than eight million Instagram posts. Some of these posts were duplicated in the dataset, as they contained more than one selected hashtag; thus these duplications had to be removed. We also deleted those users' posts, who had more than 100 posts in the corpus as we wanted to concentrate on the private users and filter out the institutional ones, who regularly posts about the same content. After the deletion of the duplications and institutional users, 4,190,031 posts remained in the initial corpus.

keywords and key phrases to look for the specific mentions that either in themselves or within their context contain those pre-given phrases that interest the user. The system gathers data almost real-time yet has a historical database that can go back up to 3,5 years. For quantitative research, data is structured by different focus points and research parameters and is visualized on an interactive dashboard. This technology also supports deep qualitative analysis as it enables thorough analysis and categorization of all the indexed web contents such as articles, posts, comments, etc., one by one.

We used two main methods: Latent Dirichlet Allocation and a neural network-based word embedding model to find the theory-relevant posts from the different contents, which remained in our corpora after the data cleaning process mentioned earlier. These methods were not initially created for the detection of content-specification, but our usage shows a possible application for this problem.

Probability-based topic models help us understand big text corpora (Liu 2015), as they are capable of exploring the hidden structure of the corpus. The Latent Dirichlet Allocation (LDA – see (Blei et al. 2003) is the most frequently used technique of topic modeling. LDA hypothesizes that words have strong semantic information, and similar documents (here, posts) use similar sets of words. Thus, LDA is a bag-of-words method (Mohr and Bogdanov 2013), where the order of the words does not matter, only their lack or existence in a given document (post) – which concept fits the use of hashtags as corpora. The concept of LDA is close to factor analysis, as its main objective is to find the latent structure that lies behind the data; however, LDA deals with texts and not with numbers. As in the case of factor analysis, each variable is assigned to each factor with some factor weight. In LDA each word is assigned to each topic with different weights. However, the algorithms behind LDA are based on a hierarchical Bayesian probabilistic model, where the topics and the distribution of topics among documents are hidden random variables, and the estimation of their conditional distribution is based on the observed collection of documents (DiMaggio et al. 2013). At the end of the process, each document (post) can be assigned to several topics with different probabilities. As LDA belongs to the unsupervised type of topic models, only the number of topics has to be given in advance, the interpretation of the final topics has to be done at the end of the process. We run the analysis in R using the *topicmodels* (Grün and Hornik 2011), *tm* (Feinerer 2018), *tidytext* (Silge and Robinson 2016), *dyplr* (Wickham et al. 2019), and *fastTextR* (Schwendinger 2017) packages.

Word embedding models can also help researchers in the understanding of large corpora. These models are called distributional semantic models as they can show the semantic representation of words in given corpora. Semantic representation, namely, the meaning of a word, is described by other words, which are the closest to a given word in a multi-dimensional (generally, a couple of hundred) vector space. Distance is originally based on the co-presence of words in a given frame, e.g., in a document or a given n-gram. We can define the width of this frame, e.g., the number of words and so decide, if we focus on the one-, two-, three-, or more words environment of a document or even on the whole document. Each type of word embedding method is based on the co-presence of words within the defined frame, and each of them results a multi-dimensional vector space. The goal is to reduce the number of dimensions (which is originally the number of words in the corpus) but preserve the distances between words. In the reduced dimensional vector space, two words will be close to each other if they share a similar environment in the corpus. Word embedding is a general term for these methods. There are different algorithms for word embedding with different advantages. One of the earliest word embedding models is the Latent Semantic Analysis (LSA), which initially took a document-word matrix created from the corpora and applied singular-value decomposition

as a data reduction technique for the creation of a lower dimensional vector space (Evans and Aceves 2016). Later models (such as word2vec, GloVe and Fasttext) apply artificial neural networks to create the vector space. They apply a minimum one hidden layer, where reduced dimensions of the vector space are the dependent variables. With these algorithms, it is possible to train bigger corpora and create higher-dimensional vector spaces (Mikolov et al. 2013). Another advantage of these newer methods is that the distribution of explained variance between the reduced dimensions is more balanced. The analysis of the corpora is realized in this vector space. Each word is represented by a vector in the vector space and distances of words are defined as cosine of the angle between vectors. For the analysis, normalization of the word vectors is important, as it means that all the words will lie on the surface of a hypersphere (same dimensions as the vector space), and the angles will be proportionate to the distances on the surface (Kozłowski et al. 2018). As word embedding methods are able to provide information about the association of words, they are useful for analyzing several problems and research questions. Some of the application of word embedding methods are about the longitudinal examination of meaning of words (Kulkarni et al. 2015) and the cultural changes (Kozłowski et al. 2018) and stereotypes (Garg et al. 2018) behind them; but these models can also identify hard-to-detect styles like sarcasm (Joshi et al. 2016). Nevertheless, word embedding methods can be useful also for the detection of word clusters, which can help to discover the latent meanings and topics of a corpus. We trained the word embedding model in R using the text2vec (Selivanov and Wang 2018) package.

As our goal is not only to show how depression- and suicide-related sociological trends appear in the digital space of Instagram but also to present, how to retrieve the theory relevant content from the wide range of suicide-related posts, some of the researcher-related decisions of the abovementioned methods will be detailed in the next chapter.

4 Ways to Find and Analyze the Relevant Content

4.1 *The Application of Topic Model*

As described in Data and Methods, altogether, 14 hashtags were used to collect the posts for our corpus, which contained more than four million posts. We applied the unsupervised topic model Latent Dirichlet Allocation, to identify the explicitly suicide-related posts, and also the main discourses of this topic. Using LDA presents two main challenges, which have to be addressed in order to fit the model.

One of the challenges of LDA is the *selection of the number of topics*. The goal is to find the right number of topics and to make sure that the chosen solution is stable. There are several ways to specify the number of topics. We chose the ldatuning package of R, which calculates different metrics to estimate the most preferable number of topics for an LDA model. We selected 5 percent random samples (200,000 posts per sample) from the original corpus and repeated this

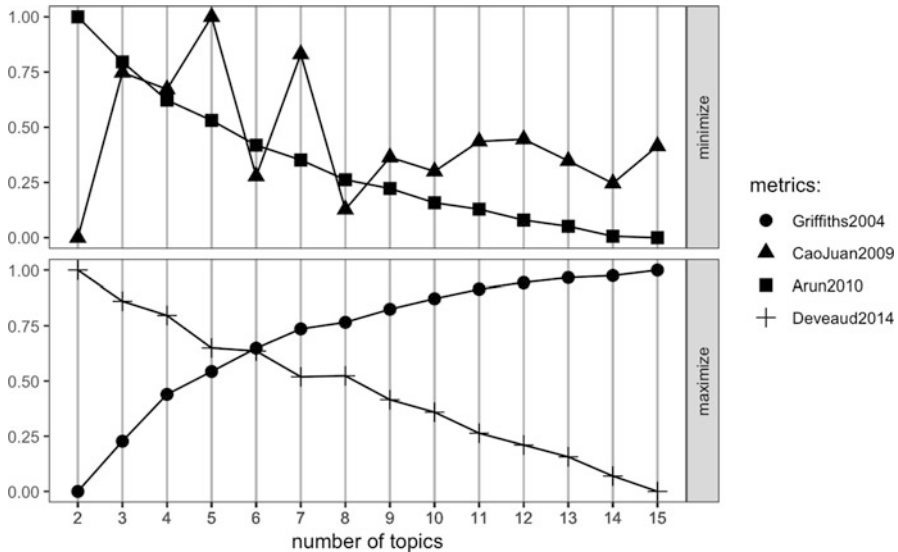


Fig. 1 Estimating the ideal topic number through different metrics

sampling process 5 times to check the stability of the structure. The 5 percent sample approach was chosen to minimize the run time of the models. On each sample, we run the `ldatuning` package with parameters from 2 to 15 topics. We did not want to go above 15 topics, because we thought we would lose the interpretability of results. Based on the results of the different models, we decided to use the Gibbs sampling method for the estimation of LDA, as the alternative VEM algorithm provided fast, but more unstable solutions. Figure 1 presents one of the results of the five samples. The results of the five samples did not show high deviations from each other.

We can interpret the different metrics plotted in Fig. 1 as the following: in the case of the metrics of Cao-Juan2009 (Cao et al. 2009) and Arun2010 (Arun et al. 2010), low values indicate good solution; in the case of Griffiths2004 (Griffiths and Steyvers 2004) and Deveaud2014 (Deveaud et al. 2014), high value indicates the good choice. Based on this figure, there is no ultimate solution, but extracting at least six topics is a reasonable choice.

In order to decide on the final number of topics, we ran several LDA models on the 5 percent samples, with 6, 8, and 10 topics. We wanted to test the stability of the results with different topic numbers. Just as in the traditional method of K-means clustering, LDA only finds local optimums; thus different runs give different results. Accordingly, to *find the most stable solution for the number of topics*, we ran two models with the same number of topics at one time and repeated this process several times. Then, we calculated the ratio of similarly classified topics for each pair of runs. The result of this calculation was around the average of 65 percent for the 6 topics approach, 60 percent for the 8 topics approach, and the ratio was much lower

for higher topic numbers. Thus, regarding stability, lower topic numbers provided a better fit.

However, overall stability was far from perfect (65 percent), even in the case of the six topic approach. Therefore, we ran the LDA 5 times, but now on the full corpus, with the six topics approach. We found that less than 50 percent of the posts were assigned to the same six topics in all of the five runs. Beyond the six base topics, there were some quite big topics, which also seemed coherent and interpretable. In the end, we decided to keep all the extracted topics, which contained at least two percent of the posts. Below 2 percent, only really small topics remained. To choose the approach of 13 topics instead of the originally selected 6 topics may seem odd; nevertheless, it was the direct result of the LDA algorithm. Like the K-means cluster, LDA also tries to avoid finding small topics. For example, if we change the number of topics from 6 to 10, LDA will not split the original six topics to smaller ones but re-run the whole analysis and tries to find 10 topics with approximately similar sizes. This is why higher topic numbers are more unstable.

To evaluate our approach, we calculated a *topic coherence* statistic, the UMass coherence measure (Mimno et al. 2011). The values of UMass are always negative, and the closer the value is to zero, the more coherent the topic is. The average topic coherence was -1.69 , but there was a high variance between the topics: from -0.81 to -3.61 . The average topic coherence improved compared to the original 6 topics models (where the value of UMass was around -2) and was similar to the simple 13 topics solution (where it was -1.7). Therefore, the 13 topics solution appeared to be an acceptable choice.

Though the number of topics was chosen, around 30 percent of the *posts were not stably classified* at this stage. In order to classify these texts, we applied the word embedding-based supervised classifier of fastText (Joulin et al. 2016).⁴ We used all stably classified posts, as a training dataset and predicted the topics of the not stably classified posts. The average value of topic coherence decreased to -1.84 , mostly because of one quite incoherent topic with a value of -3.45 .

Taking a closer look at this topic, it did not contain any relevant depression- and suicide-related content, only posts about kitchen and grooming. Although #cutting is widely used in depression- and suicide-related posts, it is used in other domains as well, for example when people post about cutting their hair. As this topic contained no relevant information from the perspective of our research question, we decided to omit it from further analysis.⁵ After the omission of this topic, the coherence value increased to -1.7 .

The preliminary analysis also showed that two topics had very similar content about mental illnesses. The list of the most important keywords is also overlapping

⁴Just for testing the accuracy of the fastText classifier, first we selected 400,000 random posts from the ones, which were stably classified by the LDA model as a training dataset. Then, we selected another 100,000 random posts from the stably classified posts as a test dataset and predicted the topic classification. The accuracy was 95 percent.

⁵The topic contained 3.5 percent of the posts of our whole sample, thus it was a relatively small one.

Table 1 Most frequent 15 words of topics of the LDA analysis

| Topic name | Words |
|---|---|
| 1. Self-harm (15,8%) | Depression, sad, depressed, suicide, anxiety, suicidal, alone, broken, sadness, selfharm, cutting, lonely, love, worthless, anorexia |
| 2. Memes (13%) | Memes, depression, meme, dankmemes, dank, funny, anime, edgy, lol, depressed, suicide, cringe, lmao, autism, filthyfrank |
| 3. Art, photography (14,1%) | Art, lonely, broken, photography, depression, mentalhealth, love, anxiety, nature, artist, drawing, instagood, photooftheday, sad, travel |
| 4. Fitness, diet (9,8%) | Mentalhealth, fitness, cutting, gym, health, motivation, bodybuilding, fitfam, anxiety, workout, fit, healthy, depression, weightloss, nutrition |
| 5. Mental health awareness (8,4%) | Mentalhealth, love, anxiety, selflove, depression, motivation, selfcare, inspiration, mindfulness, meditation, health, life, happiness, healing, positivevibes |
| 6. Illness and mental consequences (7,4%) | Anxiety, depression, ptsd, mentalhealth, stress, chronicpain, chronicillness, health, cbd, spoonie, slime, asmr, fibromyalgia, cannabis, satisfying |
| 7. Mental illness, recovery (10,1%) | Mentalhealth, anxiety, depression, mentalhealthawareness, mentalillness, ptsd, recovery, bipolar, love, therapy, mentalhealthmatters, suicide, bpd, support, stress |
| 8. Grunge, emo (6,6%) | Depression, grunge, depressed, aesthetic, sad, emo, tumblr, love, alternative, anxiety, dark, goth, music, suicide, scene |
| 9. LikeforLike (7,9%) | Love, broken, sad, lonely, quotes, depression, follow, life, like4like, depressed, heart, like, happy, likeforlike, followforfollow |
| 10. Poetry (3,3%) | Love, poetry, broken, quotes, writersofinstagram, depression, poem, writer, writing, poetrycommunity, poetsinstagram, life, poet, words, sad |
| 11. Religion (3,6%) | Love, depression, life, mentalhealth, anxiety, broken, quotes, hope, faith, god, peace, truth, motivation, jesus, inspiration |

(like #mentalhealthmatters, #recovery, #mentalhealthawareness). Based on the values of topic coherence, it seemed reasonable to merge these topics: coherence values of the original topics were -1.74 and -1.45 and -1.67 for the merged one. Based on the coherence statistics and the similar keywords, we decided to merge the two topics.

The appearance of fairly similar topics also highlighted that it would not be worth creating more topics. Ultimately, we chose an 11 topics solution. Table 1 presents the most frequent 15 words of each topic in descending order.

Topic 1 is the darkest one. It shows high probability for all the hashtags we used for the selection of the posts and contains additional disturbing words like helpme and selfhate. This is the largest topic, and it contains around 16 percent of the posts.

Topic 2 is also large, but the posts, which are assigned to it, have completely different content. This topic covers the posts about memes and meme sites like filthyfrank. It also contains those posts, which use a wide range of hashtags to reach a broader audience. Topic 9 is similar to this one as it also contains “clickbait” posts, with hashtags, such as likeforlike and followforfollow. Nevertheless, topic 9 is more about heartbreak and loneliness.

Table 2 The ratio of positive and negative words in the different topics

| | Positive | Negative |
|---------------------------------|----------|----------|
| Self-harm | 3,5% | 30,1% |
| Memes | 4,3% | 10,2% |
| Art, photography | 8,5% | 8,0% |
| Fitness, diet | 7,7% | 4,3% |
| Mental health awareness | 16,7% | 5,0% |
| Illness and mental consequences | 5,0% | 8,0% |
| Mental illness, recovery | 7,9% | 9,8% |
| Grunge, Emo | 8,1% | 12,7% |
| LikeforLike | 11,1% | 10,5% |
| Poetry | 10,0% | 12,9% |
| Religion | 18,0% | 10,1% |

Topic 3 is about photos and depression related art. This is the second largest topic, with more than 14 percent of the posts. Topic 10 (which is much smaller, only 3 percent of the posts) is similar to topic 3, as it is also connected to art, but while topic 3 is more about the visual representation, topic 10 is more about literary content. We can also include topic 8 in the art-related topics, as its posts are about specific dark music, such as grunge or goth, but hashtags like #emo or #alternative also appear here.

Topic 4 contains posts about fitness, gym, and diet associating mental health with exercising and doing sports, and their positive effect on mental health and depression. Topic 5 is also about positive messages for dealing with mental health problems, but with a focus on spirituality and positive emotions. Topic 11 covers similar perspectives, but from the point of religiosity and faith, emphasizing God and Jesus.

Topic 7 covers posts with more medical expression about mental illness than other topics. However, this topic is not exclusively negative, as it contains expressions like recovery or support as well. Medical expressions also appear in topic 6; however, it focuses on mental health issues as the consequence of other types of diseases.

To confirm our statement, that topic 1 is the darkest of all; we calculated the *sentiment value of each topic* (see Table 2 for the results). For this analysis, we used a sentiment dictionary called NRC Word Emotion Association Lexicon (Mohammad and Turney 2013). As we expected, the self-harm topic contained the most negative words: 30 percent of the words assigned to this topic were negative, and only 3.5 percent were positive. Also, this topic contained the most words of any kind of emotion (either positive or negative). Some topics, such as fitness, mental health awareness, or religion, contained more positive words than negative ones, which is also a clear sign that significantly different drivers can be identified behind the posts related to mental health or depression.

We also examined the *average number of likes and comments* of each topic (see appendix Table 3). It is not surprising that topic 9 – likeforlike – attracts the most likes and comments. However, the median value of likes is only 7, which is much

lower compared to topic 4, where the median value is 12. Thus, we can conclude that a small number of posts with a large number of likes and comments in topic 9 cause this difference. Overall, topic 1 (self-harm) has the smallest number of likes and comments (and it is true even if we use the median instead of the average). It may indicate that people respond less actively to these types of contents. However, we cannot investigate this presumption, as we do not have information about the number of followers of users who posted about raw selfharm and selfhate.

The *temporal dynamics of posts* also presents interesting results. As we do not know the number of all Instagram posts created in the different periods of the data collection, we could only calculate the relative frequency of the different topics compared to all collected posts. Thus, the number of posts in a topic was compared to all downloaded mental health-related posts, which definitely makes our analysis limited. In this section, we primarily focus on topic 1, which covers self-harm posts because we have long-time temporal trends about the number of committed suicides, which seems independent from geographical factors. Thus, if we select the topic, which is the closest to the sociological and theoretical definition of our object of interest, we can compare these trends with the ones we found on our new type of data. Figure 2 presents the yearly, monthly, and days of week trends of suicide- and depression-related posts.

A general increase can be observed in the ratio of this topic during the examined period. In 2016, 12 percent of all the collected posts belonged to this topic, which increased to 15 percent in 2017 and above 17 percent in 2018. Thus, within this wide, mental health-related space, posts about suicide, and self-harm ideation increased significantly in the examined 3 years.

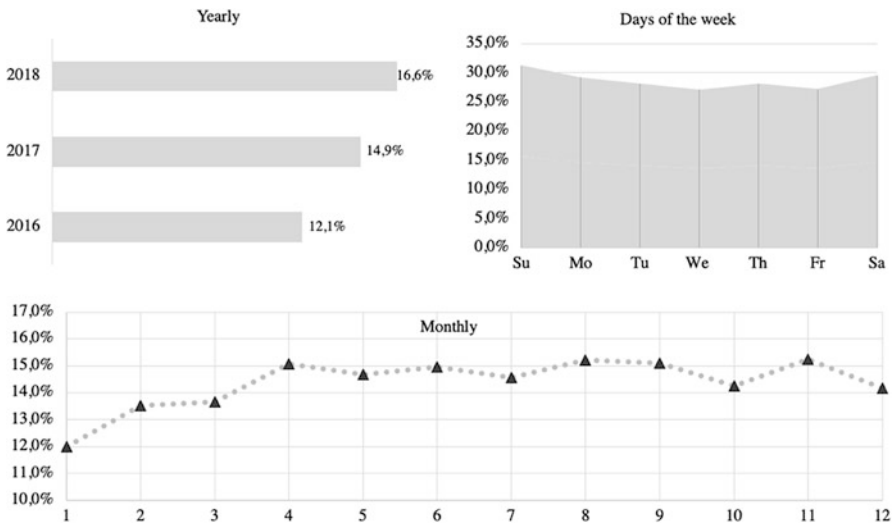


Fig. 2 Temporal dynamics of self-harm related posts (ratio compared to all posts in the dataset)

As a monthly trend, we observed that in the first 3 months of the year, the proportion of topic 1 is around 12–13 percent, then it increases to 15 percent and remains at that level.

There is another long-standing and spatially independent trend about the days of the week; thus, it is worth examining the weekly dynamic of the ratio of topic 1 among all the collected posts. The proportion of topic 1 is the highest on the weekends, and the lowest in the middle of the week (Wednesday).

Finally, we examined the ratio of topic 1 on some special days, like Christmas, New Year's Eve, and Valentine's day, just as (Beauchamp et al. 2014; Jessen et al. 1999; Zonda et al. 2009) examined these special days. However, we did not find any remarkable pattern in our data.

4.2 *The Application of Word-Embedding*

Our second approach for discovering the different discourses behind mental health-related Instagram posts is based on an artificial neural network word embedding model. We used the Glove algorithm (Pennington et al. 2014), with a 30-length window size, which means that the environment of 30 number of words was taken into account when pro-occurrences of words were calculated. This window size is larger than usual, but we wanted to make sure that all the hashtags that appeared in a post are assigned to the same window. The reason for this is that we hypothesized that the distance of two hashtags within the same post does not matter, only that they appear together in the same post. With this size of a window, each post served as one window. For the training of the vector space, we only used those words, which occurred at least 30 times in the corpus. Words with low occurrence could increase the instability of the word embedding models. Overall, more than 150,000 unique words were used in the training of the model. The trained vector space had 300 dimensions.

After we trained the model, we selected the 14 hashtags we applied in the data collection stage of the project and detected the closest words in the vector space for each of them. We used the threshold of 0.3 cosine similarity in the selection of the closest words; thus there were some initial hashtags, where less than 50 words were detected. The 0.3 value was based on our previous experiences on these word embedding models. Finally, we had 275 words in the close environment of the initial hashtags, from which we created a similarity matrix based on cosine similarities. To identify latent topics behind the corpus, we built a hierarchal cluster based on this similarity matrix.⁶ As we wanted to group the words, and we could calculate the distance between the words, clustering seemed to be a reasonable solution for this task. Based on the interpretation of the dendrogram of the hierarchical cluster

⁶As the method and logic of clustering are based on distances of the elements, we used the complement of the similarity matrix, namely dissimilarity matrix.

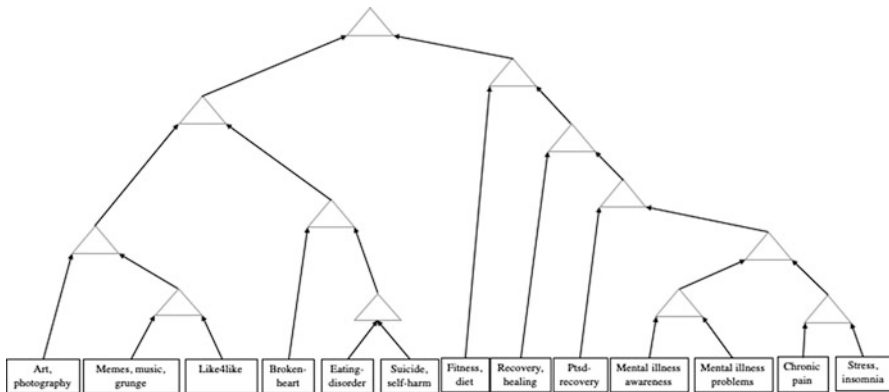


Fig. 3 Dendrogram of the hierarchical cluster of words in the vector space of the word embedding model

analysis, we decided to keep the solution of 13 word clusters. Here, the selection of the number of clusters was also not straightforward, but we could rely on the figure of the dendrogram. Figure 3 presents the simplified version of this dendrogram and thus show the distances of clusters with some interpretative labels of those. The list of all words assigned to each cluster is available in Table 4 in the appendix.

As expected, the word cluster of *chronic pain* and *stress/insomnia* are close to each other, as is *mental illness* and *mental illness awareness*. These four clusters merge before *ptsd recovery* joins them (right side of the figure). Words associated with *recovery and healing* create their own cluster, but on a higher level it merges to the first five. In the other part of the dendrogram, *suicide and self-harm*-related hashtags are close to *eating disorder*, and these two clusters are close to the cluster labeled as *broken heart*. Additionally, the cluster of *memes and music (grunge)* is close to the *likeforlike* and *followforfollow* clusters, which are also not that far from the *art and photography*-related cluster (left side of the figure). Last, but not least, we can find a distinct cluster about *fitness and diet*. This cluster is close to the recovery and mental illness-related clusters.

The main question is how similar are the results of word embedding clustering to the topics we got from the topic modeling approach. Most of the topics of the topic model can be identified in the results of the word embedding clustering, but there are some slight differences. The two main differences are the following. In the topic model, religion-related words shape a distinct topic; however, in the word embedding cluster, this topic does not appear separately. At the same time, eating disorder forms a separate cluster in the word embedding analysis, which does not appear in the topic model. Nevertheless, these are rather small differences; on the whole, the results of the two methodological approaches are exceedingly similar.

5 Discussion

People post/share/create billions of contents in social media every day, providing valuable data for researchers. However, social media data is always noisy, and it is hard to identify content germane to a particular research project. This paper had two main goals, a methodological and an analytical one.

Our methodological objective was to present a novel way to find the research related content in a large digital corpus, like the one we used from Instagram data. The combination of topic modeling and neural network word embedding methods seems to be a promising way to realize this goal. LDA topic modeling has been widely used in text mining research for a long time. However, the instability of the model and the problems regarding the selection of the correct topic number always create some uncertainty about the generalizability of the results. To solve this problem, we used a multi-step robust method to find the correct number of topics and validated the results by analysis of a word embedding model. In our approach, this neural network-based word embedding model played a limited but important role. First, it helped us to validate the results of the topic model. Second, the word embedding approach also discovered a deeper understanding of the main semantic patterns behind the posts we analyzed. The order of the application of topic modeling and word embedding is not straightforward. Reversely to our solution, there are examples in earlier researches, where word embedding results used as an input for topic modeling (see, e.g., Moody 2016). Utilized in either order, the combination of the two model families could give us more robust, more interpretable, and more valid results, which are all important for social science research.

Our analytical aim was the analysis of digital footprints of depression and suicide, on a platform, which is widely used by youngsters – Instagram. The results of the topic model showed that only 16 percent of the examined Instagram posts are connected to real suicidal ideation and/or depression. (For similar results on Twitter data, see O’Dea et al. 2015). People use depression and mental health-related hashtags in very different ways. Some people post about their (mental) illness, but it is more common to post about health (see also Muralidhara and Paul 2018), recovery, and mental health awareness (including diet and fitness). Many posts can be linked to art and poetry but also to music and music styles, such as grunge or black rock. In addition, plenty of clickbait posts and posts about memes are also present in the digital space of depression and suicide.

After identifying posts most connected to our research interest of suicide and self-harm, we analyzed the temporal dynamics of these contents. We found that it a striking result that the ratio of the self-harm-related posts increased heavily in the examined period. We do not have information about the socio-demographic composition of the users; thus, it is not possible to separate the effects, when the

same user group starts to share more negative content; and when new users, who represent a different population with more negative feelings to share appear in the digital space. As the ratio of committed suicide is higher among older people, it is also possible that what we observe here is a process, where the overall age composition of the Instagram users becomes older in the analyzed time period, resulting in the changing content of the posts. Nevertheless, as the algorithm, based on which the API of Instagram provides data, is not transparent, change in this algorithm can also be a cause of these trends.

As we do not have information about the geolocation of the data, we do not have exact population numbers about the committed suicides in this time period, to which we could compare the temporal distribution of the posts. However, in the last few years, an increasing tendency was observed in the number of committed suicides in English-speaking countries, such as the UK and the USA. In the UK, female suicide rates in 2015 reached their highest point of the decade.⁷ Overall, the suicide rates also increased in the USA, especially in the young age cohorts. Suicide rate was 10 per 100,000 inhabitants within the 15–24 age group in 2008, and 13.15 per 100,000 inhabitants within the same age group in 2016, which presents a 30 percent increase in this age cohort.⁸ Hence, it is possible that the temporal trend we detected in the digital space mirrors the same phenomenon that appears in population data.

Regarding the monthly pattern, classically analyzed trends of committed suicides show that suicides are more frequently happen during the spring and the summer (Massing and Angermeyer 1985; Zonda et al. 2005) and that these trends stand independently from the geographical location. Durkheim (1897, 2002), in his classic work, explained this trend with social factors. He observed an increase in those months (and within day periods) when social life is more active. Thus, we assumed that self-harm-related posts (topic 1) are also relatively more common during the spring and/or summer months. However, our results do not confirm this hypothesis. We could not capture the same temporal dynamic that we can observe in the case of committed suicides in the monthly trend of the posts related to self-harm. Although in January and February, the proportion of this topic (compared to all depression- and suicide-related posts) was lower than the average, in the rest of the year, there was no specific volatility between months. Thus, we could not capture the same temporal dynamics in social media as we can observe in the case of committed suicides. This confirms previous results (Kmetty et al. 2017), where the authors drew a similar conclusion analyzing Twitter data.

Weekly trends of committed suicides are also geographically and culturally independent as they show the same temporal dynamic in very different countries and continents. These trends show that suicide more frequently happens on Mondays

⁷Suicides in the UK: 2015 registrations: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/suicidesintheunitedkingdom/2015registrations> (date of last access: 10th of April, 2019).

⁸American Foundation for Suicide Prevention: Suicide statistics: <https://afsp.org/about-suicide/suicide-statistics/> (date of last access: 10th of April, 2019).

and Tuesdays and less often during the weekends. We did not find similar trends in our data; in fact, we observed an inverse tendency. The proportion of the topic (compared to all depression- and suicide-related posts) is the highest on the weekends, and the lowest in the middle of the week (Wednesday).

Our results reinforce the idea that posting about suicide is not equivalent to committing suicide. An intuitive explanation about the different temporal trends in the population and the social media can be, that there is a temporal shift between posting about- and committing suicide, however, a paper by Huang et al. (2017) does not support this explanation. They found that people who committed suicide shared the same amount of posts every day in their last month. Despite these results, we cannot say there is no relationship between suicide ideation in the social media and real-word commission of suicide. However, we can say that the link is not unconditionally direct. It is important to mention that social media companies continuously monitor these contents shared on their platform. This activity has two different effects on our data. If social media companies remove some posts or ban users (e.g., because they encourage suicide or self-harm), the composition of contents and the temporal dynamic of the posts can change. As the protocols of content moderation on the different platforms are not entirely transparent, we cannot control for this intervention. Instagram and Facebook use both algorithms and human operators to monitor the content of their platform and give signals when problematic contents appear (Chancellor et al. 2016a). Moreover, moderation has another consequence. Users, who would like to posts that will likely be moderated by the platform, start to use different expressions and words (Chancellor et al. 2016b; Moreno et al. 2016). It means that as they start using a coded language, it will be harder to capture these posts by researchers and even harder by platform mediators.

Although Instagram is primarily a photo and video sharing social media platform, we only concentrated on the texts and hashtags of the comments. A possible extension of this study may include the analysis of photos and videos, for example, with automated picture tagging techniques. Such research may have special importance, as earlier research, which dealt with the analysis of photos and videos of Instagram (Moreno et al. 2016; Reece and Danforth 2017; Brown et al. 2018), only analyzed a smaller number of cases and not large-scale data.

Wider application of our results would also be possible. Our study showed that in recent years, a mass amount of suicide- and depression-related content appeared on Instagram. Our research can provide added value in the detection of possible self-harm events, as we presented that with the utilization of complex techniques (such as topic modeling and word embedding methods), it is possible to identify the most problematic posts and most vulnerable users.

Our results also provide important input for further research on this topic. In the era of Big Data, based on the digital footprints of people, a massive amount of information is available for analysis. However, such data is very noisy and was not specifically made for scientific analyses. Thus, we have to understand the data generation mechanisms very well and also have knowledge on how to dig the valuable and useful information out from this data ocean. Our results can provide a starting point for researchers working with suicide-related social media data and can also help social scientists in the application of text analytics and machine-learning techniques.

Our paper is a good example of how the interpretative power of Social Science can be improved by incorporating the methods of Computational Social Science. It is overall really hard to research why people commit suicide. Only half of those who commit suicide have any mental illness. Every method/data which brings us closer to understand the ideation of suicide is really valuable in this research field. Moreover, social media is a perfect platform to understand people's thinking. It is also important that we did here a non-interventional study, which increases the internal validity of the results.

Appendix

Table 3 Number of likes and comments in the different topics of LDA analysis

| | Likes | Comments |
|---------------------------------|-------|----------|
| Self-harm | 15,13 | 0,62 |
| Memes | 18,04 | 0,75 |
| Art, photography | 27,78 | 1,01 |
| Fitness, diet | 27,34 | 1,35 |
| Mental health awareness | 20,87 | 0,97 |
| Illness and mental consequences | 28,91 | 1,52 |
| Mental illness, recovery | 23,10 | 1,19 |
| Grunge, Emo | 19,06 | 0,71 |
| LikeforLike | 41,16 | 1,76 |
| Poetry | 18,50 | 0,85 |
| Religion | 23,69 | 0,91 |

Table 4 Words assigned to different clusters of the word embedding model

| Cluster name | Words |
|--------------------------|--|
| Suicide, self-harm | #suicide, #suicidal, #death, #selfharm, #dead, #die, #help, #killme, #depressionquotes, #fuck, #kill, #blood, #rip, #kms, #selfhate, #helpme, #suicidegirls, #cut, #drugs, #selfharmmm, #worthless, #suicidalthoughts, #depressing, #ugly, #scars, #depressedquotes, #hopeless, #triggerwarning, #useless, #deadinside, #fat, #sadquotes, #dying, #cuts, #depressive, #numb, #cutter, #sadgirl, #cutting |
| Brokenheart | #depressed, #sad, #sadness, #alone, #hate, #cry, #hurt, #pain, #crying, #lost, #empty, #unhappy, #quotes, #tired, #lonely, #mood, #quote, #broken, #heartbroken, #tears, #feelings, #brokenheart, #heartbreak, #loneliness, #upset, #thoughts, #sadedits, #heart, #broke, #poetry |
| Stress, insomnia | #depression, #anxiety, #stress, #insomnia, #sleep, #anxietyrelief, #calm, #stressed, #relax |
| Memes, music, grunge | #emo, #memes, #meme, #funny, #music, #gay, #anime, #grunge, #tumblr, #aesthetic, #single |
| like4like | #love, #life, #girl, #happy, #me, #fear, #follow, #happiness, #smile, #friends, #like4like, #followforfollow, #l4l, #family, #instagood, #Repost, #mental, #emotional, #beautiful, #instagram, #photooftheday, #picoftheday, #people, #memories, #like, #crazy |
| Art, photography | #dark, #black, #darkness, #art, #blackandwhite, #night, #photography, #sky, #photo, #nature, #sunset, #cold, #travel, #winter, #blue, #sun, #evening, #rain, #clouds, #light, #shattered, #red, #old, #glass, #damaged, #hair, #design, #style, #haircut |
| Mental illness-awareness | #suicideprevention, #suicideawareness, #mentalhealthmatters, #endthestigma, #mentalillnessawareness, #mentalhealthrecovery, #stigma, #mentalhealthwarrior, #stopthestigma, #mentalhealthadvocate |
| Recovery, healing | #mentalillness, #mentalhealth, #mentalhealthawareness, #recovery, #hope, #selflove, #selfcare, #health, #healing, #therapy, #support, #loveyourself, #wellness, #positivevibes, #positivity, #peace, #mindfulness, #inspiration, #awareness, #motivation, #wellbeing, #healthy, #strength, #lifestyle, #healthylifestyle, #meditation, #psychology, #mind, #change, #yoga, #growth, #goals, #strong, #mindset, #mondaymotivation, #addiction, #progress, #work, #hardwork, #warrior, #freedom, #medication |
| Eating disorder | #anorexia, #anorexic, #bulimia, #ana, #eatingdisorder, #mia, #ed |
| Mental illness | #bipolar, #bpd, #anxious, #panicattack, #ocd, #socialanxiety, #adhd, #bipolarisorder, #borderlinepersonalitydisorder, #schizophrenia, #anxietydisorder, #manic, #mania, #borderline, #manicdepression, #psychosis, #bipolardepression |
| Ptsd recovery | #ptsd, #survivor, #ptsdawareness, #ptsdrecovery, #trauma, #cptsd, #veteran, #veterans, #military, #posttraumaticstressdisorder, #army, #ptsdsucks, #abuse, #PTSD, #tbi, #ptsdsurvivor |
| Fitness, diet | #fitness, #gym, #bodybuilding, #diet, #bulking, #workout, #gains, #abs, #fit, #shredded, #gymlife, #muscle, #training, #cardio, #fitfam, #food, #shredding, #protein, #fitspo, #fitnessmotivation, #gainz, #legday, #lifting, #gymrat, #dieting, #shred, #bodybuilder, #weightloss, #arms, #muscles, #physique |
| Chronic pain | #chronicpain, #chronicillness, #fibromyalgia, #invisibleillness, #spoonie |

References

- N. Andalibi, P. Ozturk, A. Forte, Sensitive self-disclosures, responses, and social support on Instagram: the case of #depression, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (ACM, New York, 2017), pp. 1485–1500. <https://doi.org/10.1145/2998181.2998243>
- R. Arun, V. Suresh, C.E.V. Madhavan, M.N.N. Murthy, M.J. Zaki, J. Xu Yu, et al., On finding the natural number of topics with latent Dirichlet allocation: some observations, in *Advances in Knowledge Discovery and Data Mining* (Springer, Berlin/Heidelberg, 2010), pp. 391–402. https://doi.org/10.1007/978-3-642-13657-3_43
- G.A. Beauchamp, M.L. Ho, S. Yin, Variation in suicide occurrence by day and during major American holidays. *J. Emerg. Med.* **46**(6), 776–781 (2014). <https://doi.org/10.1016/j.jemermed.2013.09.023>
- D.M. Blei, A.Y. Ng, M.I. Jordan, Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**(4–5), 993–1022 (2003)
- S.R. Braithwaite, C. Giraud-Carrier, J. West, M.D. Barnes, C.L. Hanson, Validating machine learning algorithms for Twitter data against established measures of suicidality. *JMIR Ment. Health* **3**(2) (2016). <https://doi.org/10.2196/mental.4822>
- R.C. Brown, T. Fischer, A.D. Goldwich, F. Keller, R. Young, P.L. Plener, #cutting: non-suicidal self-injury (NSSI) on Instagram. *Psychol. Med.* **48**(2), 337–346 (2018). <https://doi.org/10.1017/S0033291717001751>
- J. Cao, T. Xia, J. Li, Y. Zhang, S. Tang, A density-based method for adaptive LDA model selection. *Neurocomputing* **72**(7), 1775–1781 (2009). <https://doi.org/10.1016/j.neucom.2008.06.011>
- S. Chancellor, Z. Lin (Jerry), M. De Choudhury, “This post will just get taken down”: characterizing removed pro-eating disorder social media content, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (ACM, New York, 2016a), pp. 1157–1162. <https://doi.org/10.1145/2858036.2858248>
- S. Chancellor, J.A. Pater, T.A. Clear, E. Gilbert, M. De Choudhury, #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities, in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing – CSCW’16*. Presented at the 19th ACM Conference (ACM Press, San Francisco, 2016b), pp. 1199–1211. <https://doi.org/10.1145/2818048.2819963>
- Q. Cheng, T.M. Li, C.-L. Kwok, T. Zhu, P.S. Yip, Assessing suicide risk and emotional distress in Chinese social media: a text mining and machine learning study. *J. Med. Internet Res.* **19**(7) (2017). <https://doi.org/10.2196/jmir.7276>
- M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, M. Kumar, Discovering shifts to suicidal ideation from mental health content in social media, in *In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, (2016, May), pp. 2098–2110
- R. Deveaud, E. Sanjuan, P. Bellot, Accurate and effective latent concept modeling for ad hoc information retrieval. *Revue des Sciences et Technologies de l’Information – Série Document Numérique* 61–84 (2014). <https://doi.org/10.3166/DN.17.1.61-84>
- P. DiMaggio, M. Nag, D. Blei, Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics* **41**(6), 570–606 (2013). <https://doi.org/10.1016/j.poetic.2013.08.004>
- E. Durkheim, *Suicide: A Study in Sociology*, 2nd edn. (Routledge, London, 2002) [E. Durkheim (1897). *Le Suicide. Etude De Sociologie*, published by Ancienne Librairie Germer Bailliere et Cte, Felix Alcan, Editeur, Paris]
- J.A. Evans, P. Aceves, Machine translation: mining text for social theory. *Annu. Rev. Sociol.* **42**(1), 21–50 (2016). <https://doi.org/10.1146/annurev-soc-081715-074206>
- I. Feinerer, Introduction to the tm Package Text Mining in R (2018). <https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf>

- N. Garg, L. Schiebinger, D. Jurafsky, J. Zou, Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proc. Natl. Acad. Sci.* **115**(16), E3635–E3644 (2018). <https://doi.org/10.1073/pnas.1720347115>
- T.L. Griffiths, M. Steyvers, Finding scientific topics. *Proc. Natl. Acad. Sci.* **101**(suppl 1), 5228 (2004). <https://doi.org/10.1073/pnas.0307752101>
- B. Grün, K. Hornik, Topicmodels: an R package for fitting topic models. *J. Stat. Softw.* **40**(1), 1–30 (2011). <https://doi.org/10.18637/jss.v040.i13>
- X. Huang, L. Xing, J.R. Brubaker, M.J. Paul, Exploring timelines of confirmed suicide incidents through social media. Presented at the 2017 IEEE international conference on healthcare informatics (ICHI) (IEEE, 2017), pp. 470–477
- G. Jessen, B.F. Jensen, E. Arensman, U. Bille-Brahe, P. Crepet, D. De Leo, et al., Attempted suicide and major public holidays in Europe: findings from the WHO/EURO multicentre study on Parasuicide. *Acta Psychiatr. Scand.* **99**(6), 412–418 (1999)
- A. Joshi, V. Tripathi, K. Patel, P. Bhattacharyya, M. Carman, Are Word Embedding-Based Features Useful for Sarcasm Detection? (2016). arXiv:1610.00883 [cs]. <http://arxiv.org/abs/1610.00883>. Accessed 15 Apr 2019
- A. Joulin, E. Grave, P. Bojanowski, T. Mikolov, Bag of Tricks for Efficient Text Classification (2016). arXiv:1607.01759 [cs]. <http://arxiv.org/abs/1607.01759>. Accessed 16 Apr 2019
- Z. Kmetty, J. Koltai, E. Bokányi, K. Bozsonyi, Seasonality pattern of suicides in the US – a comparative analysis of a Twitter based bad-mood index and committed suicides. *Intersect. East Eur. J. Soc. Polit.* **3**(1), 56–75 (2017). <https://doi.org/10.17356/ieejsp.v3i1.302>
- A.C. Kozlowski, M. Taddy, J.A. Evans, The Geometry of Culture: Analyzing Meaning Through Word Embeddings (2018). arXiv:1803.09288 [cs]. <http://arxiv.org/abs/1803.09288>. Accessed 30 Oct 2018
- V. Kulkarni, R. Al-Rfou, B. Perozzi, S. Skiena, Statistically significant detection of linguistic change, in *Proceedings of the 24th International Conference on World Wide Web – WWW’15*. Presented at the 24th International Conference (ACM Press, Florence, 2015), pp. 625–635. <https://doi.org/10.1145/2736277.2741627>
- B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (Cambridge University Press, Cambridge, 2015). <https://doi.org/10.1017/CBO9781139084789>
- W. Massing, M.C. Angermeyer, The monthly and weekly distribution of suicide. *Soc. Sci. Med.* **21**(4), 433–441 (1985)
- T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient Estimation of Word Representations in Vector Space (2013). arXiv:1301.3781 [cs]. <http://arxiv.org/abs/1301.3781>. Accessed 15 Apr 2019
- D. Mimno, H. Wallach, E. Talley, M. Leenders, A. McCallum, Optimizing semantic coherence in topic models, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Association for Computational Linguistics, 2011), pp. 262–272
- S.M. Mohammad, P.D. Turney, Crowdsourcing a word-emotion association lexicon. *Comput. Intell.* **29**(3), 436–465 (2013)
- J.W. Mohr, P. Bogdanov, Introduction—topic models: what they are and why they matter. *Poetics* **41**(6), 545–569 (2013). <https://doi.org/10.1016/j.poetic.2013.10.001>
- C.E. Moody, Mixing Dirichlet Topic Models and Word embeddings to Make lda2vec (2016). arXiv:1605.02019 [cs]. <http://arxiv.org/abs/1605.02019>. Accessed 16 Apr 2019
- M.A. Moreno, A. Ton, E. Selkie, Y. Evans, Secret society 123: understanding the language of self-harm on Instagram. *J. Adolesc. Health* **58**(1), 78–84 (2016). <https://doi.org/10.1016/j.jadohealth.2015.09.015>
- S. Muralidhara, M.J. Paul, #healthy selfies: exploration of health topics on Instagram. *JMIR Public Health Surveill.* **4**(2), e10150 (2018). <https://doi.org/10.2196/10150>
- B. O’Dea, S. Wan, P.J. Batterham, A.L. CEAR, C. Paris, H. Christensen, Detecting suicidality on Twitter. *Internet Interv.* **2**(2), 183–188 (2015). <https://doi.org/10.1016/j.invent.2015.03.005>
- J. Pennington, R. Socher, C. Manning, Glove: global vectors for word representation, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, Doha, 2014), pp. 1532–1543. <https://doi.org/10.3115/v1/D14-1162>

- A.G. Reece, C.M. Danforth, Instagram photos reveal predictive markers of depression. *EPJ Data Sci.* **6**(1), 15 (2017). <https://doi.org/10.1140/epjds/s13688-017-0110-z>
- F. Schwendinger, *Package 'fastTextR.'* en (2017). <https://cran.r-project.org/web/packages/fastTextR/fastTextR.pdf>
- D. Selivanov, Q. Wang, *Package 'text2vec.'* (2018). <http://text2vec.org>
- J. Silge, D. Robinson, Tidytext: text mining and analysis using tidy data principles in R. *J. Open Source Softw.* **1**(3), 37 (2016). <https://doi.org/10.21105/joss.00037>
- H. Sueki, The association of suicide-related Twitter use with suicidal behaviour: a cross-sectional study of young internet users in Japan. *J. Affect. Disord.* **170**, 155–160 (2015). <https://doi.org/10.1016/j.jad.2014.08.047>
- H. Wickham, R. François, L. Henry, K. Müller, RStudio. *Package 'dplyr.'* en (2019). <http://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>
- T. Zonda, K. Bozsonyi, E. Veres, D. Lester, M. Frank, The impact of holidays on suicide in Hungary. *OMEGA J. Death Dying* **58**(2), 153–162 (2009). <https://doi.org/10.2190/OM.58.2.e>
- T. Zonda, K. Bozsonyi, E. Veres, Seasonal fluctuation of suicide in Hungary between 1970–2000. *Arch. Suicide Res.* **9**(1), 77–85 (2005)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

