What Social Media Can Tell Us About Opioid Addicts: Twitter Data Case Analysis

Completed Research

Tareq Nasralah
Dakota State University
tareq.nasralah@dsu.edu

Omar El-Gayar
Dakota State University
omar.el-gayar@dsu.edu

Yong Wang
Dakota State University
yong.wang@dsu.edu

Abstract

Opioid addiction is one of the largest and deadliest epidemics in the United States. This research investigates opioids' epidemic by analyzing recent tweets data for users who are addicted or have been addicted to opioids. Automatically analyzing social media users' posts of opioids addicted users using machine learning tools can help understand the themes and topics that exist in the up-to-date discussions of online users of social media networks. Through the analysis period from 01/01/2015 to 02/25/2019, we were able to identify 571 self-identified Twitter users. We collected a total of 20,609 English-language tweets that belong to the self-identified users. Overall, we identify the different recovery approaches, illicit drug use and user seeking for help. This study helps elicit how the daily posts of online social media users can provide a better understanding of the opioid crisis, and strengthen the public health data reporting and collection for opioids epidemic.

Keywords

Social media, opioid addict, text mining, supervised learning.

Introduction

Opioid addiction has become one of the largest and deadliest epidemics in the United States. Opioids are a group of drugs which include the illegal drug heroin and powerful pain relievers by legal prescription, such as morphine and oxycodone (Fan et al. 2017). Increased prescription of opioid medications led to widespread misuse of both prescription and non-prescription opioids before it became clear that these medications could indeed be highly addictive (Affairs (ASPA) 2017a). The U.S. Department of Health and Human Services (HHS) declared a public health emergency and consider this epidemic as a national crisis where the HHS reported that more than 130 people a day die from opioid-related drug overdoses (Affairs (ASPA) 2017b).

Social media as a big data source can be used to recognize communication and behavioral themes of problematic use of prescription drugs (Kim et al. 2017). Social media is considered as a promising and viable source of data for gaining insights of various disease conditions, patients' attitudes and behaviors, medications, etc. For example, social media can serve as a conduit to health behavioral change through messaging (Korda and Itani 2013). Despite the fact that the confidentiality and privacy of patient data are protected by the Health Insurance Portability and Accountability Act (HIPAA), social media is considered a viable source of data about patients who would willingly discuss and share health-related information about their condition. Coupled with text mining and machine learning, social media can serve as a rich resource for healthcare providers (Dredze 2012; Dredze et al. 2014).

The daily use of social media provides new opportunities for analyzing several aspects of communication. For example, social media data can be analyzed to gain insights into issues, trends, influential actors and other kinds of information (Stieglitz et al. 2018). In the Information Systems (IS) field, social media data was analyzed to investigate questions such as the influence of the social network position on information diffusion (Susarla, Oh, and Tan 2012).

Social media has been used in several studies as a resource for monitoring prescription medication abuse (Kalyanam et al., 2017; Kim et al., 2017; Sarker et al., 2016). Some of these studies' analyses showed that clear signals of medication abuse can be drawn from social media posts (Sarker et al. 2016).

In this context, this study addresses the opioid epidemic phenomenon from the opioids addicted users' perspective by collecting and analyzing recent social media posts for opioids addicts who self-identified as addicted or have been addicted to one of the opioid prescriptions. This study analyzes such data using machine learning techniques to understand the prevalent themes and perceptions that exist in the opioid addicted users' posts.

Using machine learning tools to automatically classify a large dataset of tweets that belongs to users who are addicted or have been addicted to opioids can help in understanding the nature of their issues of misusing or overdosing opioid prescriptions in addition to understanding their experiences. This can help in identifying their concerns and the common issues that they have.

In this research, we develop a strategy to identify opioid addicts over the social media, and we study the perceptions of opioid addicted users. The key implication for practice, first, it can help understand the common concerns of opioid-addicted users. Second, it can address the most discussed topics on social media for the opioid addicted users. Third, it can help get insights about the daily lifestyle of opioid addicted users by analyzing their daily posts on social media to help provide better opioid prevention, treatment, and recovery strategies. Finally, it can strengthen the public health data reporting and collection through the reports of the collected data and analysis results.

The remainder of the paper is organized as follows: the next section provides a literature review followed by the research design and methodology including data collection and analysis. The results and discussion section summarize the findings. The paper concludes with a summary of findings and contributions, and a discussion of limitations and future research.

Literature Review

Social media is used by patients to exchange information and discuss different health-related topics (Tapi Nzali et al. 2017). Online communities and social media are growing rapidly and providing new avenues for collecting evidence for policy-making processes. Popular social media platforms, including Twitter, enable new channels for their users to share information and their experiences (Zhan et al. 2017). These platforms have provided efficient methods of information access for health surveillance and social intelligence (Wang et al. 2007).

Twitter is a microblogging service where users tweet short text messages that often contain links to news stories and comments (Lerman, 2010). Several studies have used Twitter as a source of input data to identify the public's reactions to the opioid epidemic by detecting the most popular topics tweeted by users (Glowacki, Glowacki, and Wilcox 2017), for marijuana content analysis, keywords have been used to filter marijuana related tweets (Daniulaityte et al., 2015; Tian, Lagisetty, & Li, 2016) or tweets related to potential drug effects (Jiang and Zheng 2013). Other researchers studied themes describing the consequences of using marijuana by examining the related content on social media and the use of marijuana for particular situations such as Post-Traumatic Stress Disorder (PTSD) (Cavazos-Rehg et al. 2016; Dai and Hao 2017).

Several studies have addressed the drug abuse and opioid addiction. Here we present a summary of these articles, techniques and methodologies used, and reported results. Following a systematic review, Maglione et al., (2018) have studied the effects of medication-assisted treatment for opioid use disorder on functional outcomes. The analysis showed that weaknesses in the body of evidence prevent strong conclusions about the effects of medication-assisted treatment for opioid use disorder on functional outcomes.

Kalyanam, Katsuki, R.G. Lanckriet, & Mackey, (2017) developed a strategy in the field of digital epidemiology to better identify, analyze, and understand trends in non-medical use of medications and drugs prescription. Tweets were filtered for three commonly abused drugs, namely Percocet, OxyContin, and Oxycodone. The primary themes identified evidence of high levels of social media discussion about polydrug abuse on Twitter.

Shyamashree et al. (2017) examined electronic health record data to study the trends of prescription opioid dependence using observational database with a natural language processing-based NoSQL architecture. The objective was to help prioritize interventions in vulnerable population subgroups. Results showed that the predominance is among the Non-Hispanic, white population in the 19 to 38 years of age group. The prevalence in younger age implies that the complications related to opioid dependence would become a costly burden of disease for a longer duration of time.

Fan et al. (2017) have designed a framework to automatically detect the opioid addicts from Twitter. Tweets were collected using a crawler based on keywords related to opioids, such as heroin and morphine, as well as users' profiles. Then the meta-path-based approach is used to formulate similarity measures over users and different similarities are aggregated using Laplacian scores. The results showed that knowledge from daily-life social media data mining could support a better practice of opioid addiction prevention and treatment.

Cherian, Westbrook, Ramo, & Sarkar (2018) have characterized representations of codeine misuse through analysis of public posts on Instagram to understand text phrases related to misuse. A total of 1156 sequential Instagram posts, related to opioid medication and text phrases associated with codeine misuse, were analyzed using content analysis to identify common themes arising in images, as well as culture around misuse, including how misuse is happening and being perpetuated through social media. Results showed that codeine misuse was commonly represented with the ingestion of alcohol, cannabis, and benzodiazepines.

Social media users' posts were used to better understand providers' attitude toward using recovery drugs such as 'naloxone' to treat opioid addictions (Haug et al. 2016). Indeed, social media, such as Twitter, can help in building frameworks to automatically detect the opioid addicts and support a better practice of opioid addiction, prevention, and treatment (Fan et al. 2017)

Several studies have used social media as a source of input data to identify individuals amenable to drug recovery interventions (Eshleman, Jha, and Singh 2017) and use text mining to examine and compare discussion topics of social media communities to discover the thematic similarity, difference, and membership in online mental health communities (Park, Conway, and Chen 2018).

Glowacki, Glowacki, & Wilcox (2017) have utilized text mining to analyze the public's reactions to the opioid crisis. The authors identified the public's reactions by identifying the most popular topics tweeted by users. A total of 73,235 original tweets and retweets were collected over two months. The tweets collected depend on limited keywords related to opioids, all tweets contained references to "opioids," "turnthetide," or similar keywords. Tweets were analyzed to identify the most prevalent topics using topic modeling.

The aforementioned research affirms the potential for analyzing users' posts on social media as a mechanism to better understand their needs and perceptions toward drug addiction and more specific opioid prescription medication abuse. In this research we emphasize the population of the study to the users who self-identified as they are addicted to opioids or have been addicted to opioids. We expand the focus to study the users experience by including different terms such as opioid drug abuse terms related to symptoms, Rx drugs, illicit drugs, slang drug terms, related activities and behaviors, recovery, pre- and post- conditions, location and time period.

Research Design and Methodology

To study opioid users' experience and addiction, we leveraged Crimson Hexagon, a social media analytics for data collection and analysis (Hexagon n.d.). Crimson Hexagon (CH), a social media analytics company, employing an unsupervised and supervised machine learning techniques and text analysis model developed by Daniel Hopkins and Gary King (Hopkins and King 2010). In general, the steps proceed as follows: First, data collection: This is where the user determines the date range of interest, the social media data sources, the keywords to use to search for relevant posts, and the restrictions to impose (language, geographic location...etc.) Second, analytics: This is where the user selects the appropriate approach (using predefined categories for sentiment or opinion analysis, or user-defined categories). For the latter, the user identifies pertinent categories and a labeled data set for training purposes.

Data Collection

The data for this study collected from Twitter, where we collect tweets from users who self-identified as they are addicted to opioids or have been addicted to opioids in all over the United States. Table 1 shows the opioid users search query.

To retrieve the relevant users' tweets, we define different terms that related to the opioid drug abuse topic keywords by considering the drug abuse terminology such as topic terms related to symptoms, Rx drugs, illicit drugs, slang drug terms, related activities and behaviors, recovery, pre- and post- conditions, location and period. The collected tweets are all selected based on the criteria of having at least one related keyword, and we excluded retweets and addresses as shown in Table 2.

```
(("I am addicted" OR "I was addicted" OR "I am addict" OR "I addict" OR "I addicted" OR "I have been addicted")
 AND
 (Opioid OR Opioids OR Opiates OR Opiate
 OR Naloxone, Propoxyphene OR Hydrocodone Vicodin OR oxycodone OR Oxycontin OR Oxy OR Oxys OR Percocet OR Oxymorphone OR
Opana OR Morphine OR Hydrocodone OR Tramadol OR Fentanyl
 OR Duragesic OR Actiq OR Subsys)) ~2
  AND -

(http OR https OR RT)
```

Table 1. Search query for the opioid addicted users

```
(
Opioid OR Opioids OR Opiate OR Opiates
   OR Heroin OR Kratom OR Marijuana OR Hashish OR Weed OR Opium OR cannabis OR Cocaine OR Crack
   OR Codeine OR Naloxone OR Propoxyphene OR Hydrocodone OR Vicodin OR Oxycodone OR OxyContin OR Oxy
   OR Oxys OR Percocet OR Oxymorphone OR Opana OR Morphine OR Hydromorphone OR Tramadol OR Fentanyl
   OR Duragesic OR Actiq OR Subsys OR Recovery_Drugs OR Methadone OR Dolophine OR Methadose OR Diskets
   OR Naltrexone OR Revia OR Vivitrol OR Buprenorphine OR Probuphine OR Subutex OR Suboxone
   OR
    ((
    (addict OR pain OR overdose OR overdoses OR high OR cough OR misuse OR
     pharmacy OR pharma OR Friend OR Friends OR Dealer OR doctor)
    OR
    (self-medication OR pain OR "severe pain" OR withdrawal OR high OR cough OR surgery OR
    intranasal OR smoking OR injection OR plugging OR oral OR snort OR sniff OR
    milligram OR bags OR pills OR pill OR millilitre OR bottles OR bottle)) AND (Opioid OR Opioids OR Opiate OR Opiates
    OR Heroin OR Kratom OR Marijuana OR Hashish OR Weed OR Opium OR cannabis OR Cocaine OR Crack))
    OR #opioid OR #kratom OR #opioidcrisis OR #chronicpain OR #fentanyl OR #OpioidEpidemic OR #overdose
    OR #opioidhysteria OR #iamkratom OR #opioids
)
AND -
(http OR https OR RT)
```

Table 2. Search query for the addicted users' tweets

Data Analysis

Crimson Hexagon employs the ReadMe algorithm developed by Daniel Hopkins and Gary King (Hopkins and King 2010). This is a supervised learning algorithm that expects the researcher to hand-code a 'training set" of documents (posts) into a set of predefined categories. Crimson Hexagon provides an already 'trained' model for sentiment and opinion mining, or an opportunity for the researcher to train their own model using user-defined categories.

The ReadMe algorithm is particularly suited when the objective is to know the proportion of the population of posts that fit in specific categories. Rather than calculating this proportions based on the categorization of individual posts, ReadMe gives approximately unbiased estimates of category proportions even when the optimal classifier performs poorly (Hopkins and King 2010).

The key advantage of using a social media analytics platform such as Crimson Hexagon is that it provides access to the "Twitter fire hose", i.e., it provides access to every public tweet ever posted on Twitter in any language and from any geographic location that meets the search criteria. While it provides the possibility of downloading data for further analysis and exploration, a limitation of Crimson Hexagon is the constraints imposed (mostly by Twitter) on the amount of data the researcher can download. We have addressed this limitation by manually reading and verifying thousands of tweets.

In this research, we use the ReadMe (provided by Crimson Hexagon) to analyze the proportion of tweets that fall into specific categories. We initially utilized Crimson's 'built-in' categories and associated 'trained' algorithm to explore the general opinion surrounding the opioid addiction and use.

We trained a model to identify the proportion of tweets falling into customized. The categories are primarily drawn from the literature pertaining to the opioid epidemic. Most notably following Fan et al. (2017), we identified taking illicit drugs (such as heroin), using Medication-Assisted Treatment (MAT) and seeking for help. In addition, we included using cannabis as alternative to recover, identified in Glowacki et al. (2017) and added a number of categories that are related to other approaches to recover and getting opioids. Examples include using kratom to recover from opioids, trading opioids, taking other illicit drugs and needing opioids. Appendix 1 describes each of the categories, keywords delineating each of these categories, and a representative tweet. Using Appendix 1 as a code book, we manually labeled and distributed 320 tweets over the 10 categories. The training was an iterative process ensuring that each category is clearly outlined by the examples. The number of the coded tweets increased over several runs of the model as we reviewed the categories and coded more tweets. The performance of the model improved in classifying the tweets in alignment with the predefined categories for the labeled data set.

Results and discussion

Over the period from 01/01/2015 to 02/25/2019, only 571 self-identified Twitter users were retrieved. Overall, 291 (51% of all users) of the users included gender information while 61 (11% of all users) included age information distributed as shown in Figure 1. We collected a total of 20,609 English-language tweets using the search query shown in Table 2 for the self-identified opioid addicted users over the same period, from 01/01/2015 to 02/25/2019.

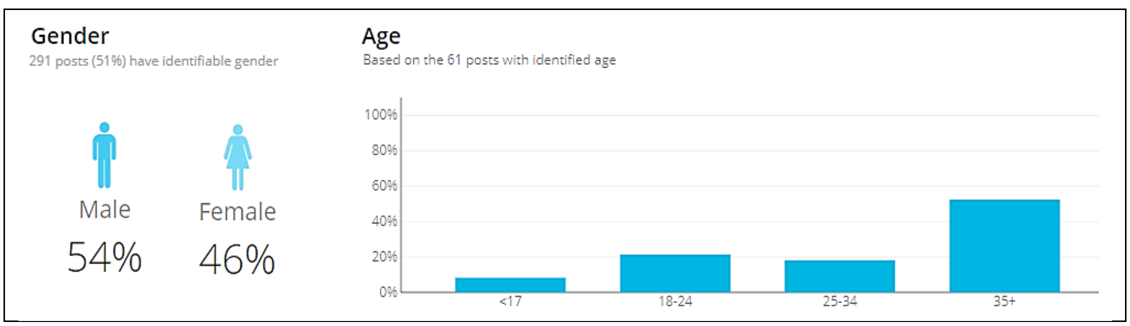
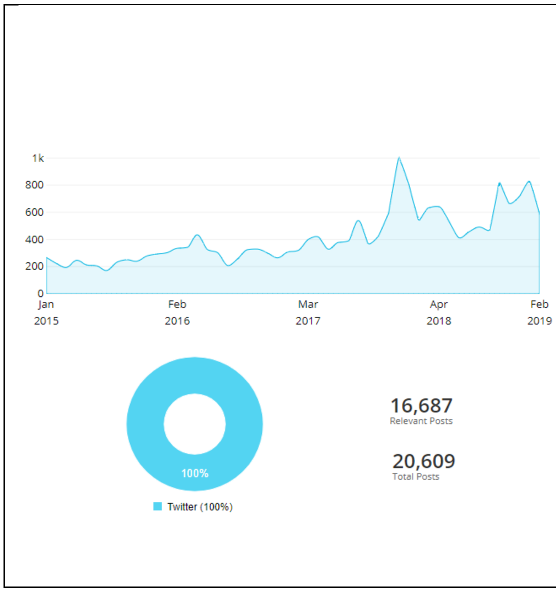


Figure 1. Users demographic information (Gender & Age)

Figure 2 showing the volume of the collected tweets. In the figure there is several peaks of the users posts. Those peaks happen when there is some news is going on over the media about the opioid epidemic. For example, some peaks happen when there is news about a famous character who facing an issue with opioid drugs, or if there is a news about an announcement from the U.S. government about the opioid epidemic, as the figure 2 shows, there is a peak around the last quarter of the year 2017. In that time, acting Health and Human Services (HHS) Secretary issued a statement upon declaring a nationwide public health emergency regarding the opioid crisis. Some of the tweets have no related content with the study context, for example:" *when I get a kitten I'm naming it morphine*", this tweet classified as irrelevant. After excluded the irrelevant posts we end up with 16, 687 posts.

Figure 3 is a summary of the proportion of tweets falling into the various categories and the percentage of the total relevant tweets for each category. Overall, the results demonstrate the identified categories account for 81% (16,687 out of 20,609) of the total number of posts. Obviously, the results show five main categories: "In Recovery" (38%), "Taking illicit drugs" (27%), "Seeking for help" (20%), "Trading Opioids" (12%), and "Needing Opioids" (3%). For the "In Recovery" category we were able to define some subcategories related to the approach of the addicted user follow to recovery. These subcategories are, "Using Medication-

Assisted Treatments (MAT)" (18%), "Using Cannabis" (15%), and "Using Kratom" (5%). Also, under "Taking Illicit drugs" category, we have subcategories for the types of the illicit drugs that addicts users take, these subcategories are, "Cannabis" (14%), "Cocaine" (9%), and "Heroin" (4%).



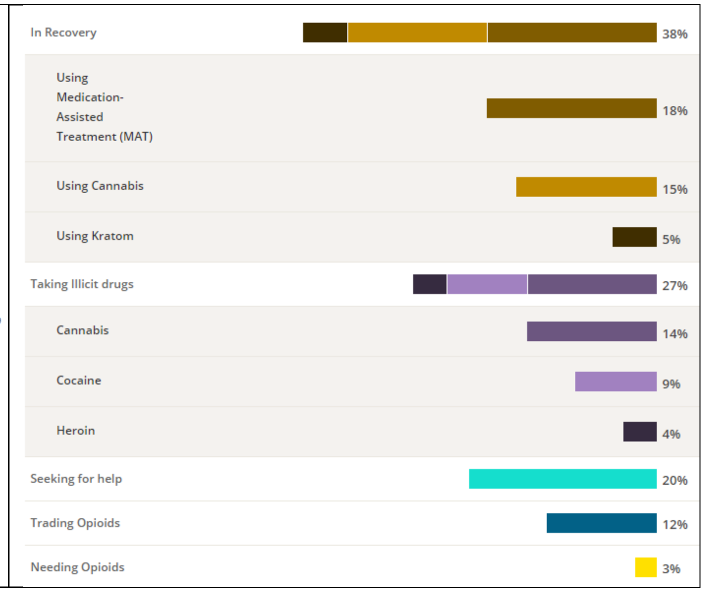


Figure 2. Data volume over the period **Figure 3. Proportion of tweets by category**

There are a number of posts that are particularly prevalent, most notably, users in recovery, taking illicit drugs and seeking for help. In essence, posts pertaining to recovery (what the Opioid's users have used to recover from their opioid's addictions and overdose) accounted for 38% of the tweet volume with 18% related to using of the Medication-Assisted Treatment (MAT). Tweets related to taking illicit drugs (namely, Cannabis, Cocaine and Heroin) amounted to 27% of the tweet volume while the seeking for help category accounted for 20%. Trading opioids amounted for 12%, while needing opioids (tweets for users who are looking to get opioids drugs) was responsible for a mere 3%. Figure 4 provides a high-level view of keyword clusters and their relations using a sample of 1,000 tweets. Overall, two clusters relate to the use of illicit drugs such as, heroin, cocaine and weed, and using kratom and cannabis for recovery.

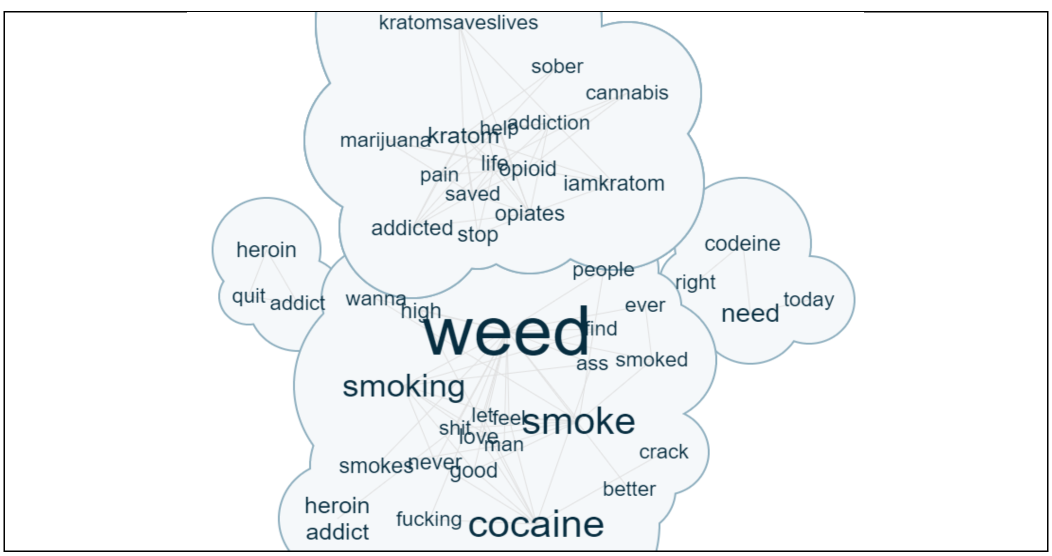


Figure 4. Cluster of keywords from 1000 tweets of all categories

Overall, the identified categories in the analysis results show the different recovery approaches that the opioid addicted users take to manage their misuse and addiction. Also, the other illicit drugs that they are taking, and they may addict to. Also, the results show the users need for help and information during their health management. The results highlight different areas where opioid addicted users may need some sort of interventions and innovations strategies using social media the healthcare provider and policy makers can formulate to assets in facing the opioid addiction.

Addressing the most discussed topics on social media that related to drug abuse, such as opioids epidemic, can help understand the problem dimensions and create the proper strategies. Examples of such strategies can be getting insights from the discussion topics to make the opioid media campaigns more effective in preventing opioid misuse, as well, addressing the most important topics can help in telling how the opioid addicted users' opinions can provide a tool to improving the opioid recovery programs.

Conclusion

Online social media is a rich source to collect data about individual's daily activities, interests and lifestyle. Applying text mining techniques can help in understanding the concerns of online social communities. In this research we examine the opioids addicted users' posts to understand the recent themes and perceptions that exist in their posts. We used supervised machine learning to automatically analyze the content of the opioids addicted users' tweets.

From a theoretical perspective, this research highlights the importance of further developing and adapting text mining techniques to social media for drug abuse. Such media represents inherent challenges for text mining given the amount of noise and distortion in the data. Of particular significance is the emphasis on developing approaches for improving the discovery and identification of the drug abuse topics in social media domains characterized by a plethora of highly diverse terms and a lack of commonly available dictionary/language by the drug abuse communities such as in the opioid drug abuse case. For example, future research can aim at developing a social media text mining framework for drug abuse. The framework can have phases to address the challenges of the discovery and identification of the drug abuse topics over the social media and insuring the quality and pertinent of the collected data. Accompanying the framework is formulating a drug abuse ontology for social media that can help in providing a systematic approach to study the different drug abuse phenomena.

From a practical perspective, automatically analyzing social media users' posts using machine learning tools can help understand the users' themes and topics that exist in the up-to-date discussions of online users of social media networks. By doing so, this could help better recognize the recent status of the opioid epidemic and other drug abuse. Addressing the most discussed topics on social media that relate to drug abuse can help understand the problem dimensions and create the proper strategies. Examples of such strategies can be getting insights from the discussion topics to make the opioid media campaigns more effective in preventing opioid misuse, as well, addressing the most important topics can help in telling how opioids users' opinions can provide a tool to improving opioids recovery programs. Specifically, by identifying prevalent categories such as opioid addicted users who are in recovery using medication-assisted treatment (MAT) and users who seeking for help, interventions can aim towards supporting such individuals by ensuring sufficient providers, and providing personalized messages of encouragement to seek or continue to seek treatment. Awareness of such categories can also inform social media campaigns about the MAT programs, including opioid treatment programs (OTPs), this can help in spreading awareness among the users and help them in managing their addiction.

The research limitations and possibilities for improvement can be through additional refinement of the defined categories, and focusing on specific category, e.g., seeking for help. In addition, enhancing this research with surveys of opioids users to better understand their specific concerns and experience.

REFERENCES

Affairs (ASPA), Assistant Secretary of Public. 2017a. "What Is the U.S. Opioid Epidemic?" Text. HHS.Gov. December 4, 2017. https://www.hhs.gov/opioids/about-the-epidemic/index.html.

2017b. "HHS.Gov/Opioids: The Prescription Drug & Heroin Overdose Epidemic." Text. HHS.Gov. December 21, 2017. https://www.hhs.gov/opioids/.

Cavazos-Rehg, Patricia A., Shaina J. Sowles, Melissa J. Krauss, Vivian Agbonavbare, Richard Grucza, and Laura Bierut. 2016. “A Content Analysis of Tweets about High-Potency Marijuana.” Drug and Alcohol Dependence 166 (September): 100–108. https://doi.org/10.1016/j.drugalcdep.2016.06.034.

Cherian, Roy, Marisa Westbrook, Danielle Ramo, and Urmimala Sarkar. 2018. “Representations of Codeine Misuse on Instagram: Content Analysis.” JMIR Public Health and Surveillance 4 (1): e22. https://doi.org/10.2196/publichealth.8144.

Dai, Hongying, and Jianqiang Hao. 2017. “Mining Social Media Data on Marijuana Use for Post Traumatic Stress Disorder.” Computers in Human Behavior 70 (May): 282–90. https://doi.org/10.1016/j.chb.2016.12.064.

Dredze, Mark. 2012. “How Social Media Will Change Public Health.” IEEE Intelligent Systems 27 (4): 81–84. https://doi.org/10.1109/MIS.2012.76.

Dredze, Mark, Renyuan Cheng, Michael J. Paul, and David Broniatowski. 2014. “HealthTweets.Org: A Platform for Public Health Surveillance Using Twitter.” In Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence. https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/view/8723.

Eshleman, Ryan, Deeptanshu Jha, and Rahul Singh. 2017. “Identifying Individuals Amenable to Drug Recovery Interventions through Computational Analysis of Addiction Content in Social Media.” In , 849–54. IEEE. https://doi.org/10.1109/BIBM.2017.8217766.

Fan, Yujie, Yiming Zhang, Yanfang Ye, Xin li, and Wanhong Zheng. 2017. “Social Media for Opioid Addiction Epidemiology: Automatic Detection of Opioid Addicts from Twitter and Case Studies.” In CIKM’17, November 6-10, 2017, Singapore, 1259–67. ACM Press. https://doi.org/10.1145/3132847.3132857.

Glowacki, Elizabeth M., Joseph B. Glowacki, and Gary B. Wilcox. 2017. “A Text-Mining Analysis of the Public’s Reactions to the Opioid Crisis.” Substance Abuse, July, 1–5. https://doi.org/10.1080/08897077.2017.1356795.

Haug, Nancy A., Jennifer Bielenberg, Steven H. Linder, and Anna Lembke. 2016. “Assessment of Provider Attitudes toward #naloxone on Twitter.” Substance Abuse 37 (1): 35–41. https://doi.org/10.1080/08897077.2015.1129390.

Hexagon, Crimson. n.d. “AI-Powered Consumer Insights Company | Crimson Hexagon.” Accessed June 14, 2018. https://www.crimsonhexagon.com/.

Hopkins, Daniel J., and Gary King. 2010. “A Method of Automated Nonparametric Content Analysis for Social Science.” American Journal of Political Science 54 (1): 229–47. https://doi.org/10.1111/j.1540-5907.2009.00428.x.

Jiang, Keyuan, and Yujing Zheng. 2013. “Mining Twitter Data for Potential Drug Effects.” In Advanced Data Mining and Applications, 434–43. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-53914-5_37.

Kalyanam, Janani, Takeo Katsuki, Gert R.G. Lanckriet, and Tim K. Mackey. 2017. “Exploring Trends of Nonmedical Use of Prescription Drugs and Polydrug Abuse in the Twittersphere Using Unsupervised Machine Learning.” Addictive Behaviors 65 (February): 289–95. https://doi.org/10.1016/j.addbeh.2016.08.019.

Kim, Sunny Jung, Lisa A Marsch, Jeffrey T Hancock, and Amarendra K Das. 2017. “Scaling Up Research on Drug Abuse and Addiction Through Social Media Big Data.” Journal of Medical Internet Research 19 (10): e353. https://doi.org/10.2196/jmir.6426.

Korda, Holly, and Zena Itani. 2013. “Harnessing Social Media for Health Promotion and Behavior Change.” Health Promotion Practice 14 (1): 15–23. https://doi.org/10.1177/1524839911405850.

LERMAN, K. 2010. “Information Contagion : An Empirical Study of the Spread of News on Digg and Twitter Social Networks.” Proceedings of 4th International Conference on Weblogs and Social Media (ICWSM), 2010. https://ci.nii.ac.jp/naid/20000917072/.

Maglione, Margaret A., Laura Raaen, Christine Chen, Gulrez Azhar, Nima Shahidinia, Mimi Shen, Ervant Maksabedian, Roberta M. Shanman, Sydne Newberry, and Susanne Hempel. 2018. “Effects of Medication Assisted Treatment (MAT) for Opioid Use Disorder on Functional Outcomes: A Systematic Review.” Journal of Substance Abuse Treatment 89 (June): 28–51. https://doi.org/10.1016/j.jsat.2018.03.001.

Park, Albert, Mike Conway, and Annie T. Chen. 2018. “Examining Thematic Similarity, Difference, and Membership in Three Online Mental Health Communities from Reddit: A Text Mining and Visualization Approach.” Computers in Human Behavior 78 (January): 98–112. https://doi.org/10.1016/j.chb.2017.09.001.

Sarker, Abeed, Karen O'Connor, Rachel Ginn, Matthew Scotch, Karen Smith, Dan Malone, and Graciela Gonzalez. 2016. "Social Media Mining for Toxicovigilance: Automatic Monitoring of Prescription Medication Abuse from Twitter." Drug Safety 39 (3): 231–40. https://doi.org/10.1007/s40264-015-0379-4.
Shyamashree, Sinha, Burstein Gale R, Leonard Kenneth E, Murphy Timothy F, and Elkin Peter L. 2017. "Prescription Opioid Dependence in Western New York: Using Data Analytics to Find an Answer to the Opioid Epidemic." Studies in Health Technology and Informatics, 594–598. https://doi.org/10.3233/978-1-61499-830-3-594.
Stieglitz, Stefan, Milad Mirbabaie, Björn Ross, and Christoph Neuberger. 2018. "Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation." International Journal of Information Management 39 (April): 156–68. https://doi.org/10.1016/j.ijinfomgt.2017.12.002.
Susarla, Anjana, Jeong-Ha Oh, and Yong Tan. 2012. "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube." Information Systems Research 23 (1): 23–41. https://doi.org/10.1287/isre.1100.0339.
Tapi Nzali, Mike Donald, Sandra Bringay, Christian Lavergne, Caroline Mollevi, and Thomas Opitz. 2017. "What Patients Can Tell Us: Topic Analysis for Social Media on Breast Cancer." JMIR Medical Informatics 5 (3): e23. https://doi.org/10.2196/medinform.7779.
Wang, Fei-Yue, Kathleen M. Carley, Daniel Zeng, and Wenji Mao. 2007. "Social Computing: From Social Informatics to Social Intelligence." IEEE Intelligent Systems 22 (2): 79–83. https://doi.org/10.1109/MIS.2007.41.
Zhan, Yongcheng, Ruoran Liu, Qiudan Li, Scott James Leischow, and Daniel Dajun Zeng. 2017. "Identifying Topics for E-Cigarette User-Generated Contents: A Case Study From Multiple Social Media Platforms." Journal of Medical Internet Research 19 (1): e24. https://doi.org/10.2196/jmir.5780.

	Category	Description	Keywords	Examples
1	In Recovery - Using Medication-Assisted Treatment (MAT)	Highlights the different Medication-Assisted Treatment (MAT) that the Opioid's users have used to recover from their opioid's addictions and overdose.	"Saved my life", clean, sobered, recovered, "opioid free", ...etc.	I was addicted to Vicodin for years. Had jail/prison/forced rehab. None of it helped. A praying mom and Suboxone saved my life. Been clean seven yrs. & live a 'normal' life now.
2	In Recovery - Using Cannabis	Indicates the use of Cannabis by Opioid's users to recover from addiction and using opioid prescriptions drugs	Marijuana, hashish, Weed, cannabis, pot, "saved me", helped, "legalize", ... etc.	I was addicted to opioids and sleeping pills from 7-13ish because A doctor prescribed them to me. The one thing that helped me get off them was pot. Nothing else helped me sleep, eat and remove pain like that. Without the consequences of true addiction.
3	In Recovery - Using Kratom	Captures use of Kratom by Opioid's users to recover from addiction and using opioid prescriptions drugs	Kratom, sober, recovery, "kratom save lives", clean, ... etc.	#kratomsaveslives #iamkratom. I was addicted to opiates and kratom has kept me.clean for 4 years. It is a miraculous plant that saves lives.
4	Taking Illicit drugs - Cannabis	Relates to opioid's users who have using Cannabis for their pain managements or they have appreciation for it overall.	Marijuana, hashish, Weed, cannabis, pot, "medical marijuana", smoke, ... etc.	"I really like smoking Marijuana". "I like to smoke weed, it ain't a bad habit"
5	Taking Illicit drugs - Cocaine	Relates to opioid's users who have using Cocaine for their pain managements or they have appreciation for it overall.	Cocaine, Crack, alternative, legalize, ... etc.	I'm addicted to crack cocaine I'm sorry u all had to find out this way
6	Taking Illicit drugs - Heroin	Relates to users who are addicted, or they have taken heroin	Heroin, "China white", addicted, shot, ...etc.	Yeah I was a morphine addict 4 a while I shot it up then moved 2 heroin b4 I quit & just got a script for suboxone to not be the way I was
7	Seeking for help	Captures the tweets for users who are seeking help and explanations	Help, how, what, need, someone, please, ... etc.	Please Dont Let The Codeine Hit My System
8	Trading Opioids	Captures the tweets that take about buying the opioids prescriptions drugs form the illegal sources	Dealer, selling, buy, bought, opioid, someone, street, ... etc.	"Lp kids buy pill pressed fentanyl and think it's Xanax". "if you know someone selling fentanyl send them my way. I think it's almost that time."
9	Needing Opioids	Opioid users who are looking to get opioids drugs	Oxy, Vicodin, pain, Percocet, need, ...etc.	First time I've ever run out of pain meds this soon. I need Percocet/Oxy like big time. In excruciating pain right now.
10	Irrelevant (off-topic)	The post that has no related content for any of the above categories		my mom thinks literally everything smells like weed

Appendix 1. Codebook for labeling categories