Discovery of Informal Topics from Post Traumatic Stress Disorder Forums

Reilly Grant
Grinnell College
Grinnell, IA, USA
grantrei@grinnell.edu

David Kucher
University of Michigan
Ann Arbor, MI, USA
dkucher@umich.edu

Ana M. León
University Juárez Autónoma de Tabasco
Tabasco, MX
aleon@tabasco.edu

Jonathan Gemmell*
School of Computing DePaul University
Chicago, IL, USA
jgemmell@cdm.depaul.edu

Daniela Raicu
School of Computing DePaul University
Chicago, IL, USA
draicu@cdm.depaul.edu

***Abstract*—Post Traumatic Stress Disorder (PTSD) is a public health problem afflicting millions of people each year. It is especially prominent among military veterans. Understanding the language, attitudes, and topics associated with PTSD presents an important and challenging problem. Based on their expertise, mental health professionals have constructed a formal definition of PTSD. However, even the most assiduous mental health professionals can care for only a small fraction of those suffering from PTSD, limiting their perspective of the disorder.**

As social networking sites have grown in acceptance, users have begun to express personal thoughts and feelings, such as those related to PTSD. This wealth of content can be viewed as an enormous collective description of PTSD and its related issues. We automatically extract informal latent topics from thousands of social media posts in which users describe their experience with PTSD and compare these topics to the formal description generated by mental health professionals. We then explore the pattern and associations of these topics. Our evaluation demonstrates that we can automatically extract meaningful PTSD related topics and discover meaningful interactions among them.

***Keywords*—Post Traumatic Stress Disorder (PTSD), Topic Modeling, Word Embeddings, Association Rules**

I. Introduction

Post Traumatic Stress Disorder (PTSD), a psychiatric disorder that can occur after a traumatic effect such as sexual assault, warfare, or threats to a person's life, is a serious mental health problem. Symptoms of PTSD include disturbing thoughts, mental or physical distress to trauma-related prompts, recurring dreams, and an increase in the fight-or-flight response [5].

It is estimated that over 9% of Americans will experience PTSD at some point in their life, and 3.5% of Americans experience PTSD in any given year [5]. Furthermore, approximately 30% of people who have spent time in active war zones are afflicted by PTSD. For wounded veterans, this number jumps to 70% [7]. This is concerning, as PTSD is associated with increased risk of violence, unemployment, and struggles with interpersonal relationships [18].

Given the prevalence and implications of PTSD, understanding its impact on those who suffer from it poses an important public health objective. Previous work has shown that PTSD has a marked effect on the language of those afflicted, which can be indicative of how users may respond to treatment [3], [37]. This effect is particularly strong when patients are discussing their traumatic experience. This indicates that a valuable way to gain information about PTSD is to analyze the language of those afflicted by the disorder.

PTSD is currently rigorously defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM) by eight criteria required for a diagnosis [5]. These criteria include, among other things, exposure to death, threatened death, serious injury or sexual violence, intrusive thoughts or nightmares, and overly negative thoughts and assumptions about oneself or the world. While these criteria are useful for mental health professionals attempting to diagnose and treat patients, they are not well suited for automatically analyzing social media posts generated by those suffering from PTSD, as patients rarely speak in such clinical terms.

As social networks have become accepted in society, individuals regularly post about their personal lives in online forums and other social media platforms. Some of these forums are dedicated to PTSD, and thus contain a large corpus of related text. In this work, we use topic modeling algorithms to extract informal latent topics from data containing the thoughts, feelings, and expressions of those with PTSD. Topic modeling is a machine learning technique used to discover latent topics from a corpus of text. Using this tool can result in the discovery of recurring themes and topics such as the types of abuse, violence, and trauma

We perform topic modeling on over seven thousand submissions to r/PTSD, a forum focused on providing a place for people suffering from PTSD and their family members to share their experiences and seek support. First, we learn semantic embeddings for words in the submissions using a topic modeling algorithm. Then, the word embeddings are clustered to produce informal latent topics, which can be compared to the DSM description of PTSD. Finally, we produce association rules describing the ability of one topic to predict another within a post.

The results of our experiment demonstrate that latent topics related to PTSD can be automatically generated. We also found that these topics can be compared to the criteria described by the DSM for the prevalence of PTSD, but some topics were related to aspects of the individual's personal life and the impacts of the disorder. We were able to demonstrate that topics can be meaningfully analyzed by comparing them to each other using association rules, and furthermore that this

*Corresponding Author

is useful for learning about the context to learn more about the context in which words are used. These findings lead us to conclude that social media data can be used to evaluate the mental health of individuals, and using these platforms as a way to address PTSD in the population may be viable in the future.

The rest of this paper is organized as follows. Section II describes related work. In Section III, we discuss how we discover topics from social media submissions about PTSD using Word2Vec and *k*-means clustering, after which we describe how we applied association rules to the resulting topics. The forum we used as a data source is presented in Section IV. Our results, along with an analysis of these results are discussed in Section V. Finally, we conclude the paper by addressing our findings and the direction of future work in Section VI.

II. Related Work

Social media has been used as a source of data to characterize mental health by numerous researchers [10], [28], leading to the development of computational tools, such as the Linguistic Inquiry Word Count (LIWC) [39]. Researchers have attempted to predict depression, identify suicidal Twitter posts, and analyze the expression of PTSD in social platforms [13], [23], [36].

Much of the research analyzing social media posts rely on the LIWC. This widely validated tool is able to analyze text for semantic and syntactical information [39]. This tool has been used to analyze text, often in the form of social media posts, related to suicide, depression and PTSD [3], [6], [10], [11], [13], [14], [17], [23], [29]. A significant study found that when asked to describe their "trauma stories", the respondents' use of words related to death were significantly predictive of how well a patient would fare [3].

PTSD has previously been analyzed within Twitter [23]. Harman, et al., used regular expressions to identify users who self-diagnosed as suffering from PTSD, then used a combination of the LIWC and languages models to analyze the data [23]. They found that their models were able to differentiate between users that self-diagnosed as suffering from PTSD and random users. They also found that areas around military bases had a higher proportion of PTSD tweets and that more active bases had a higher proportion of tweets that had word use indicative of PTSD.

Researchers have also analyzed the social media platform Reddit, where they focused on the subreddit r/SuicideWatch. Researchers have found that the language used on Reddit differs among subreddits, which indicates that there are identifiable language patterns in some forums designated to mental health issues. [17]. This finding exposed how certain mental issues could affect the way people communicate with others, suggesting that people with certain mental health conditions could be identified by their use of language on social media. Another study attempted to infer which users would transition from discussing mental health to suicidal ideation [14]. Additionally, we previously found that topics discovered by analyzing the language used by r/SuicideWatch have been shown to have a significant relation to risk factors that experts have identified. This previous work reinforces the motivation to analyze Reddit data for informal topics pertaining to mental health.

This work proposes using computationally generated language models to analyze text surrounding PTSD. Language models include simple bag-of-word models [12] and extend to more sophisticated models such as probabilistic latent semantic analysis [25], latent Dirichlet allocation [8], and Word2Vec [34], [35]. These models have been used to explore many different topics, such as comparing discovered topics in data [26], creating recommendation systems [44], [46], and comparing different languages [45].

We focus on the Word2Vec language model developed by Mikolov et al. [34], [35]. Word2Vec attempts to create an efficient way to represent words which can accurately capture semantic features from the text. To validate the utility of their model, the researchers demonstrated that it could capture the deep and interesting semantic relationships between words. For example, the models could capture the semantic relationship between "King" and "Queen" and the relationships between countries and their capital cities. After generating a language model, we cluster the words from the model into informal topics.

We also use association rules to find patterns among the informal latent topics we discover. We apply association rule learning using a frequent pattern tree approach to find patterns in how different topics are associated with each other [1], [2], [21]. Association rules have previously been used to find relationships in financial data [22], medical diagnosis data [15], legal data [27], and sales data [4]. Also, given survey information, association rules have been used to find patterns in adolescent willingness to see a counselor [20].

We extend the efforts of previous work. First, whereas previous work in analyzing mental health topics contained in a corpus of social media text often attempts to impose a presupposed structure upon data, we instead attempt to automatically extract the structure from the data itself. Our previous work evaluating suicidal ideation showed that is possible to reproduce many of the formal topics from social media posts. Inspired by our results on informal topic discovery for suicidal ideation, in this work we discover and compare our data-driven topics to the formal definition of PTSD as described by DSM [5]. We then expand upon these efforts by discovering significant associations between topics as expressed by users.

III. Methodology

In this section, we provide a detailed description of our procedure, including how we generate vector representations of words using Word2Vec, use *k*-means clustering to produce topics in text data, and then use association rules to find patterns among our topics.

A. Word Embeddings

In our model, we represent words as vectors of real numbers [41]. More formally, each word $\vec{w}$ is represented as:

$$\vec{w} = \langle \phi(i_1), \phi(i_2) ... \phi(i_n) \rangle \quad (1)$$

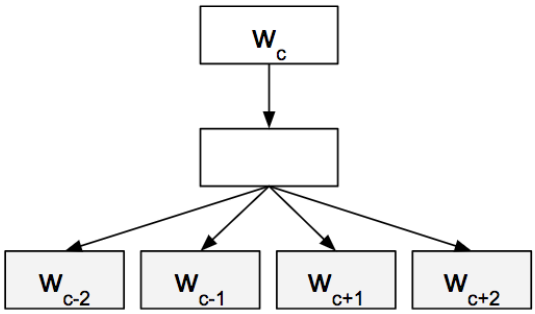

Fig. 1. Architecture for the skip-gram model. The skip-gram model predicts the distributed representations of neighbors given a word. In this figure, the representation has a window size of 2, where w_c is the target word being evaluated, and w_{c+i} denotes the surrounding context words.

where $\phi(i_1)$ through $\phi(i_n)$ represent the weights of the *i*th word in the vector space. Conceptually, these word representations fill a high-dimensional vector space, where relative word locations encode semantic information.

Methods to generate these embeddings mapping include neural networks [34], [35] dimensionality reduction [30], [32], probabilistic models, [19] and explicit representation in terms [33]. Word embeddings have been shown to improve common natural language processing tasks such as sentiment analysis [43] and syntactic parsing [42].

B. Word2Vec

In this work, we turn our attention to Word2Vec [34], [35], which has been argued to have many advantages over other topic modeling algorithms, including latent semantic indexing [38], latent Dirichlet allocation [8], and non-negative matrix factorization [31].

Word2Vec is a collection of related models used to extract word embeddings from a corpus of text. In this work, we leverage the skip-gram variant of Word2Vec, which predicts neighboring words of a target to create its vector representation. See Figure 1.

Unlike many neural network models, Word2Vec uses only a single hidden layer, making the algorithm relatively efficient. Negative sampling is used to speed up training further by updating only selected neurons in the network after every iteration. Finally, sub-sampling is employed to give more attention to rare words and less attention to common words such as "the" which add little meaning to the model and add time to training the model.

Learning the word representation is achieved by performing back-propagation on our training examples. Finally, the softmax function normalizes the output of the neural network, so that sum of all outputs is equal to 1. Similar words are next to each other in vector space because they are likely to show up in the same context. This captures rich semantic information, which allows for an accurate language model to be built. Word2Vec performs accurately when predicting text, comparing similar words, and in analogical reasoning [35].

C. Clustering

While individual word representations contain semantic information, clusters of word representations can create a meaningful grouping of related words.

Clustering algorithms group similar items together. Different algorithms have different metrics for determining similarity, and therefore create different clusters. Examples of metrics include Manhattan distance, cosine similarity, and Euclidean distance. We use Euclidean distance, as we use the relative positions of vector representations to determine similarity.

Some clustering algorithms are agglomerative clustering, DBSCAN, and *k*-means clustering. We leverage *k*-means clustering because of its ability to create simple, localized clusters [24]. This algorithm randomly places *k* cluster centers in vector space, then assigns each item, in this case, the items are vector representations of words, to the nearest cluster center. Next, the mean location of each cluster center is calculated by averaging the contents of the cluster, and the cluster center is moved to that location. This is repeated until there are no new assignments of word representations to cluster centers. The results in *k* clusters, which are represented by the words constituting them. More formally, one can look at a cluster as a vector of words, C_i:

$$\vec{C_i} = \langle w(i_1), w(i_2)...w(i_n) \rangle \tag{2}$$

where $w(i_1)$ through $w(i_n)$ represent the *n* words constituting the *i*th cluster.

Collections of words, such as our clusters of word vectors, can be viewed as topics, which are evaluated based on the words which constitute them. For example, a cluster containing the words "join", "sports", "team", "joined", "practice", and "won" describes the topic of playing team sports. Therefore, we can identify topics within a corpus by analyzing the clusters we create.

D. Association Rules

After using clustering to find informal topics in our data, we then leveraged association rule mining to explore how the topics relate to each other. An association rule is a rule of the form: $C_i \implies C_j$, where C_i and C_j each represent a cluster. These rules are found by analyzing the co-occurrence of items within a large collection of sets. For example if the data set was a list of grocery store purchases, and people who bought butter also tended to buy milk, then the association rule for this would be represented by "$butter$" $\implies$ "$milk$".

In this work, we use the frequent pattern (FP) growth algorithm for finding association rules [2], [21]. We choose this algorithm because of its efficiency. Because we are working with a large amount of data relative to other association rule mining tasks, the efficiency of the tool we use is of paramount importance. The FP-growth algorithm gains its efficiency through more compactly storing the data it is using in a unique tree called a Frequent Pattern tree. This FP-tree stores sets sorted by their frequency, in a way that allows the crucial information about the frequency of items in a data set to be represented more compactly, and accessed far more quickly than representation as a collection of sets [21]. By using this data structure, the FP-growth algorithm is capable of doing an

analysis of much large and more complicated data sets than would otherwise be possible.

To perform association rules on forum posts we first represent words in the latent feature space of their embedding. After clustering words, we are then able to represent words by the clusters to which they have been assigned. Since the posts themselves can be represented by the words they contain, it is then possible to multiply these two matrices and thus represent posts by the clusters – or informal topics – contained in each post.

Association rules allow us to discover which topics are related to each other, and what their relationship is. By finding the association rules that are present in our posts between different word clusters, we hope to find revealing patterns in how topics are connected for those suffering from PTSD.

IV. r/PTSD

Reddit, the 8th most popular website in the world according to Alexa[1], is a forum-based platform for users to have online discussions with other users about many topics. The platform contains many "subreddit", or sub-forums, which focus on particular topics which help organize the submissions. Approximately 6% of adults online have used Reddit [16].

The subreddit, r/PTSD, is a forum on Reddit dedicated to those suffering from PTSD and their families. Users are able to talk about their experiences and seek advice. Frequent topics of conversation on r/PTSD include talking about family members with PTSD, asking for advice for what to do about one's PTSD, and expressing the emotional anguish they feel as a result of their PTSD. At the time of our data collection, r/PTSD had 9,640 subscribers[2].

Because of the sensitive nature of r/PTSD, malicious users are less likely to post inflammatory comments. Moreover, the forum is heavily moderated. As a result, r/PTSD is a remarkably clean dataset, as off topic or offensive posts are quickly removed by moderators. This contrasts with much of the social media data available online.

We collected all submissions from r/PTSD's inception, in August of 2008, through November of 2016. Although on Reddit, users have the ability to leave comments on every post, we did not collect comments for this data analysis. Posts are frequently used by those with PTSD to describe their experiences, while comments on the posts are often expressing support or advice. Due to the focus of this paper, we chose to focus on the text of the original post.

Before running our analysis we extensively cleaned our data. We first removed deleted, empty posts from our data set We then substituted the word "link" for all links, and removed non-alphabetical characters such as numbers, punctuation, and special characters. Finally, the text and the title of the posts were combined, as to not omit any text from the analysis. After cleaning our data, there were 7,057 posts from 3330 users. These posts contained a total of 1,592,918 words, with 24,942 unique words. Researchers interested in the data or code for this analysis are invited to contact the authors.

[1] www.alexa.com

[2] www.reddit.com/r/PTSD

V. Results

In this section, we assess the models created with the r/PTSD data. For reproducibility, we first present the parameters used in the experiments. Next, we assess individual words to determine if they capture semantic information. These words are clustered, at which point we evaluate the quality of the clusterings in the form of topics. Then, the automatically generated topics are compared to the criteria of PTSD presented in the Diagnostic and Statistical Manual of Mental Disorders. Finally, we discuss the discovered association rules and summarize our analysis of our results.

A. Experimental Parameters

Once we obtained the data we began by creating vector representations using Word2Vec [40]. Each post was processed using the window size of 5, common in the literature, which looks at the previous and next 5 words, along with the current word, looking at a total of 11 words at once.

Negative sampling was set to 20 on the recommendation of the authors of the Word2Vec model, based on the size of our data [34], [35]. The authors recommend a range of 5-20 for smaller documents, with a larger value increasing the accuracy of the model while also increasing training time. After extensive evaluation, we chose to represent words with a vector of 300 features using the skip-gram model and hierarchical softmax. These parameters seemed to provide rich semantic descriptions while minimizing computational overhead.

In order to preserve the meaning of phrases, we turned common phrases into single tokens, called n-grams. This allows phrases such as "new york" to be separate from "new" and "york" alone, which have very different meanings. This resulted in an increase in the size of vocabulary to 28,137 unique words and phrases. We set the minimum count for a word to be included in the model at 10 occurrences. This threshold removed noise in the data such as misspelled words and unrecognized characters among other things. After filtering our vocabulary, we decreased our vocabulary to its final size of 7787 unique words.

Next, we clustered the vector representations of words by using k-means clustering [9]. An important input to the algorithm is the selection of k. We chose a value of 100 because it offered a sufficient number of clusters to capture the topics of the posts without being too large to manually evaluate. Regarding an error as the distance of each vector to its cluster center, we calculated the sum of the squared errors (SSE) for values of $K = 25, 50, 75, 100, 125, ...375, 400$. The knee of the SSE curve was approximately 100 clusters.

B. Analysis of Word Representations

To test the quality of our automatically generated word representations, we evaluated the representations of "medication", "military" and "abuse". Table I contains the test words along with the closest, or most similar, words in the vector space. For example, "medication" appears with the words most similar to it, "medications" and "meds".

For each of the three words shown in the table, the closest words share similar semantic information. The word "medication" is closest to words that are synonyms and more

specific examples. In the case of "military", the related words are closely related conceptually. Finally, "abuse" is nearest in proximity to words which capture specific types of abuse.

medication	military	abuse
medications	tours	neglect
meds	active duty	emotional abuse
prozac	army	sexual abuse
anti depressants	vietnam	physical abuse

TABLE I. NEAREST WORDS IN VECTOR SPACE TO TEST WORD.

After looking at semantic similarity, we sought to test our model with respect to analogical reasoning, as it was done by Mikolov et al. [35]. We use vector algebra on the word representations.

For example, we computed "son" - "boy" + "girl" a new vector is created that most closely matches the vector for "daughter". This type of analogical reasoning indicates that our model has captured rich semantic relationships among the words.

We also explored semantic relationships more in line to the domain of PTSD. We found that "abuse" + "young" = "molestation", and that "veteran" + "trauma" = "trauma survivor". This indicated to us that the word embeddings have captured rich semantic information relevant to the ideas surrounding PTSD.

These examples match the concepts we associate with words and their relationships, and thus we believe that the words can be clustered to extract topics.

C. Analysis of Informal Topics

In our evaluation of our automatically generated clusters, we considered the most frequent words in 100 clusters. For example, one cluster contained the words "ptsd", "symptoms", "diagnosis", and "was diagnosed". All of these words have semantic similarities, related to the concept of PTSD. Further examining this cluster, we also find "depression", "adhd" and "bipolar disorder". This overall reflects the topic of being diagnosed with a disorder in general.

An additional example of these clusters having semantic meaning is a cluster including the words "trauma" "happened" "past", and "remember". These words all relate to the concept of a past trauma. The cluster also includes "flashback", and "reliving", indicating that this cluster captures the theme of re-experiencing traumatic memories. In addition to providing evidence for these clusters capturing semantic information, this cluster also provides evidence for clustering finding topics relevant to PTSD, as major criteria for identifying PTSD is re-experiencing traumatic events [5].

Finally, we found that our clustering picked up on some topics less related to PTSD. For example, one cluster included the terms "person", "love", "happy", and "nice", while another included "read", "link", "information", and "online". Both clusters seem to have a coherent semantic relationship, but neither directly relates to the idea of PTSD.

D. Comparison to DSM criteria

In this section, we present the eight criteria for diagnosing PTSD described in the Diagnostic and Statistical Manual of Mental Disorders [5]. We then evaluate the ability of the informal topic models to capture these criteria by evaluating the common words in each cluster and comparing it to the description of the criterion. Unless otherwise noted, all clusters are presented by showing the five most frequent words in the cluster. When less common words convey more meaning we include them and mark them with an asterisk.

1) Criterion A: The person was exposed to: death, threatened death, actual or threatened serious injury, or actual or threatened sexual violence, in the following way(s): direct exposure, witnessing the trauma, learning that a relative or close friend was exposed to a trauma, or indirect exposure to aversive details of the trauma, usually in the course of professional duties (e.g., first responders, medics).

In Table II we report the clusters which correspond to this criterion by including five words from each cluster, or words that were particularly indicative of the rest of the cluster. For example, one of the selected clusters contains "police", "involved", "killed", "shooting" and "assaulted".

police*	abuse	child	against
involved*	childhood	physically	beat
killed*	rape	emotionally	in front of
shooting*	physical	violent	himself
assaulted*	sexual	abusive	gun

TABLE II. RELATED TOPICS TO CRITERION A.

All these clusters relate to the idea of a traumatic event, but they do so in different ways. The first cluster we have "killed", "shooting", and "assaulted". It seems, that this cluster reflects the idea of armed assault, which fits under the label of actual or threatened serious injury, or, actual or threatened death.

The next two clusters both reflect the trauma of abuse, although in different ways. The cluster including "abuse", "rape", and "sexual", consists primarily of descriptions of abuse, elsewhere in the cluster using words like "mental abuse" and "emotional abuse". This seems to indicate that this cluster is describing abuse in the abstract, or abuse happening to another person. In contrast, the cluster which includes "physically", and "emotionally" reflects more descriptions of having suffered from abuse. The inclusion of "growing up" in this cluster further indicates that this cluster reflects suffering from abuse for an extended period of time.

The final cluster associated with this criterion seems to reflects more the trauma of physical abuse. In addition to the word "beat", this cluster includes "grabbed", "hitting", "punched", and "belt". While the previous cluster describing physical abuse seems more to reflect the experience of it over a long period of time, this cluster reflects the description of a single episode of physical abuse.

Our analysis of this criterion, and it's associated clusters reflects the extent to which trauma is a broad category. There are both many different types of trauma and many different ways to describe these traumas.

2) Criterion B: The traumatic event is persistently re-experienced, in the following way(s): intrusive thoughts, nightmares, flashbacks, emotional distress after exposure to traumatic reminders, and physical reactivity after exposure to traumatic reminders.

We examine the topics associated with this criterion and further look at how these topics are associated with the ways in which the criteria can manifest.

flashbacks	physical reactivity
trauma	literally
remember*	shaking
memories*	starts
reliving*	tears
flash back*	scream

TABLE III. SUBTOPICS FOR CRITERION B.

Consider Table III where two clusters have been manually assigned to formal topics of "flashbacks" and "physical reactivity". We see that these clusters correspond closely with the ways in which people can re-experience traumas. The first cluster, including "trauma" and "reliving", strongly reflects the idea of having a flash back to a moment, while the second cluster very strongly reflects the idea of having a physical reaction to a memory of a trauma. However, these two clusters are relatively easy to label for this criterion. Other clusters seemed to capture multiple topics. Consider Table IV

anxiety	triggered	having
flashbacks	flashback	worse
nightmares	panic attack	having flashbacks
panic attacks	dream	having panic attacks
dreams	nightmare	having nightmares

TABLE IV. RELATED TOPICS TO CRITERION B.

Each of these clusters fits into the broader category of re-experiencing traumatic events. However, instead of distinguishing flashbacks, panic attacks, and nightmares, the informal topics have instead distinguished different syntactic variation of these ideas such as "panic attacks", "panic attack", and "having panic attacks".

This finding indicates that when expressing themselves online, those suffering from PTSD may not discuss some concepts with as much precision as mental health professionals. In these cases, syntactic differences can overwhelm some of the semantic nuances between words like "flashbacks" and "nightmares".

3) Criterion C: Avoidance of trauma-related stimuli after the trauma, in the following way(s): trauma-related thoughts or trauma-related reminders.

This criterion is a good example of the differences between the formal topics proposed by domain experts and the informal topics generated by topic modeling. While it is true that those suffering from PTSD attempt to avoid stimuli related to the trauma, it is very difficult to find evidence of this avoidance within textual data. Indeed, one would assume that if they were avoiding thinking about specific stimuli they would often not post about it online.

We thus were unable to find any clusters that fit with criterion C.

4) Criterion D: Negative thoughts or feelings that began or worsened after the trauma, in the following way(s): inability to recall key features of the trauma, overly negative thoughts and assumptions about oneself or the world, exaggerated blame of self or others for causing the trauma, negative affect, decreased interest in activities, feeling isolated, or difficulty experiencing positive affect.

This criterion, in contrast to criterion C, was particularly abundant and nuanced. The freedom to express oneself is after all one of the appeals of online forums. Consider first the symptom of having overly negative thoughts and assumptions about oneself or the world. We can see in Table V a cluster that contains the theme of being overly negative.

Overly Negative
everything
mind
hurt*
dying*
forever*

TABLE V. OVERLY NEGATIVE SUBTOPIC FOR CRITERION D.

This cluster further contained the words "ive lost" and "drowning", indicating to us that this cluster strongly captured the theme of depressive feelings.

Now, consider the symptom of exaggerated blame of self or others. We found that the cluster represented in Table VI captured the idea of blaming oneself and blaming others, and that it further contained "so guilty" and "betrayed by", gives additional evidence that this cluster captured the theme of blaming oneself and others.

feel
felt
sad
ashamed
guilty*

TABLE VI. BLAME SUBTOPIC FOR CRITERION D.

Finally, we consider the symptom of negative affect. We found three clusters associated with this topic, all listed in Table VII In each of these clusters, users are describing their emotional state.

fear	feeling	im
pain	scared	hurt
thoughts	angry	afraid
feelings	tired	hate
stress	makes me	kill myself

TABLE VII. NEGATIVE AFFECT SUBTOPICS FOR CRITERION D.

It seems that individuals expressing their experience with PTSD online often include a description of their emotional state. Moreover, this emotional state often aligns with the DSM criterion of PTSD.

5) Criterion E: Trauma-related arousal and reactivity that began or worsened after the trauma, in the following way(s): irritability or aggression, risky or destructive behavior, hypervigilance, heightened startle reaction, difficulty concentrating, or difficulty sleeping.

We were successful at finding clusters related to having trouble sleeping and irritability, and with being on high alert. Table VIII reports these clusters. The cluster we associated with trouble sleeping additionally had the words "can't sleep" and "couldn't sleep".

Trouble Sleeping and Irritability	High Alert
sleep	ball
sometimes	mode
usually	fight or flight
cry	crawl
wake up	heart rate

TABLE VIII. RELATED TOPICS FOR CRITERION E.

The first cluster describes difficulty sleeping, a common symptom associated with PTSD. The second cluster captures the heightened stress level and hypervigilance of trauma survivors including how they "ball" their fists, have increased "heart rates", and find themselves "fight or flight" "mode".

For some dimensions of this DSM criterion, we could not find corresponding clusters. For example, there is no related cluster for "destructive behavior", likely because users are less likely to admit these symptoms online.

6) Criterion F: Symptoms last for more than 1 month.

According to the DSM, in order to be diagnosed with PTSD, it is required that the symptoms have persisted for at least one month. While Word2Vec embeddings can capture rich semantic information such as the notion of "weeks" and "depressed", it struggles to capture sentence level concepts, such as "I have been depressed for weeks." Thus establishing the length of time symptoms have been occurring using raw text is currently out of the scope of this work.

now	year	after
ive	month	during
been		months
since		last
have been		half

TABLE IX. RELATED TOPICS FOR CRITERION F.

However, we do find many clusters related to time, such as those found in Table IX. Future work may attempt to parse this type of content, to separate out posts based on the time frame they are discussing.

7) Criterion G: Symptoms create distress or functional impairment (e.g., social, occupational).

Whereas in the previous criterion is was challenging to associate symptoms with time, in this criterion it is somewhat easier to associate the symptoms with particular contexts and outcomes. See Table X.

no	work
money	job
pay	school
paid*	working
bills*	college

TABLE X. RELATED TOPICS FOR CRITERION G.

The first cluster includes words such as "money", "pay", and "bills". Alone these words might suggest the cluster capture the topic of finances. However, this cluster also includes "no", "couldn't afford", and "no money" suggesting that the informal topic captured by this topic really described the individual struggle with finances.

In the second cluster, we see a similar story. The top most terms include "work" and "school". However, additional words in the cluster include "low", "unemployed", "grades", "dropping", and "failing". By looking deeper into this informal topic, we see that when individuals are discussing work and school they are often talking about their struggle with keeping a job or maintaining their grades.

Ultimately, this criterion seems to be one which is more challenging to identify through text analysis than the other clusters we have been identifying, but still reflects the criterion as defined by mental health professionals.

8) Criterion H: Symptoms are not due to medication, substance use, or other illness.

Criterion H is a requirement listed in the DSM. Assessing whether or not a symptom is due to trauma or to medication, substance use, or other illness is a challenging task for a clinician and not within the capabilities of topic modeling.

However, we are able to identify clusters related to medication, drugs and other mental disorders, such as those in Table XI.

medication	drugs	disorder
meds	alcohol	bipolar
medications	drug	causes
mg	weed	dsm
prazosin	marijuana	textbook

TABLE XI. RELATED TOPICS FOR CRITERION H.

It seems that individuals posting on the forum often talk about medications and other mental health issues. The first cluster capture the notion of over the counter medications such as "prazosin" or further into the cluster "xanaz", "zoloft", "antidepressants", "lexapro", "klonopin", "prozac", and "wellbutrin". The second cluster captures the notion of recreational drugs. The third cluster captures language pertaining to the diagnostics of PTSD and other mental health disorders.

9) Additional Informal Topics: For the eight criterion established by the DSM, we saw that in some cases the informal topics learned by clustering words correlated well. For example, the symptoms of flashbacks and hypervigilance were well represented by two of the informal topics. In other cases, criterion established by the DSM was more challenging for our topic modeling approach to capture, particularly those associated time or those where the individual might not be willing to admit the symptom. In this section we turn our attention to topics that are not described by the DSM.

home	parents	watching	car
house	mother	watch	driving
live	my mom	movie	walking
leave	father	play	ran
place	my dad	music	across

TABLE XII. INFORMAL TOPICS OF DAILY ACTIVITIES.

In Table XII we see topics related to the individuals' household, family, entertainment, and transportation. The inclusion of these topics underscores the impact of PTSD on daily life. When discussing PTSD in a casual setting, individuals often speak of their relationship with their family or how watching a movie can trigger a panic attack.

In Table XIII we see topics related to PTSD. The first captures the idea of "therapists" and types of treatment. The

therapy	treatment	military
therapist	free	combat
doctor	mental health	army
psychiatrist	medical	iraq
emdr	research	afghanistan

TABLE XIII. INFORMAL TOPICS RELATED TO PTSD.

second cluster captures the research and free treatment opportunities often posted on the forum. The final cluster clearly captures the notion of military service.

E. Association Rules

After evaluating our clusters by comparing them to the criteria given by the DSM, we now turn our attention to exploring the relationship between clusters using association rules. We converted every post into a vector, where the *i*th entry indicates the number of words from the *i*th cluster in that post. We then generated frequent itemsets of clusters from this data using a support threshold of 10%, meaning we only looked at combinations of elements that appeared in at least 10% of the posts. For example, a frequent itemset might show that the topic of "military service" and the topic of "flashbacks" occur together in more than 10% of posts. We selected association rules with a confidence higher than 50%, meaning the rule is true in at least 50% of the cases where it is applicable. We restricted the format of our rules to one to one to make analysis more tractable. Our final set of rules had 2,521 different one to one rules.

We evaluated all one to one rules manually to find the meaningful associations. Thousands of association rules were generated. We describe a few of the more significant ones here. See Table XIV.

We found that $1 \implies 2$, with a confidence of 0.79, and a lift of 5.97. Cluster 1 includes the words "medication" and "meds". Cluster 2 includes the words "ptsd" and "symptoms". The confidence value of 0.79 means that in approximately 79 percent of the posts, if cluster 1 was present, cluster 2 would also be present. Furthermore, the lift value of 5.97 indicates that when cluster 1 appeared, cluster 2 was about 6 times more likely to appear than it occurs in general. In short, this association rule states that posts including a discussion about medications also include a discussion about symptoms. Such a rule is not surprising, but it gives us confidence that the process of association rule mining has picked up significant real patterns in the data.

Another rule discovered was $3 \implies 2$ with a confidence of 0.77 and a lift of 4.62. This rule allows for better understanding of cluster 3 in the context of PTSD. Before knowing about this rule, cluster 3 seems to be simply about sounds. However, after learning about this association, we can see that in PTSD forums sounds are often described in the context of symptoms. For example, certain sounds might invoke depression or other symptoms associated with the disorder.

We also see that there is an association rule between cluster 3 and cluster 4 with a confidence of 0.66 and a lift of 3.94. This rule reveals another insight into cluster 3; sounds can trigger PTSD symptoms. Such an interpretation goes beyond the semantic meaning often captured by word embeddings and begins to capture a larger narrative of the users' experience with PTSD.

In another rule we found that $5 \implies 6$ with a confidence of 0.74 and a lift of 3.59. The context of cluster 5 is used in is associated with "trauma" and "memories" and thus cluster 5 is often used when posters are recalling their traumas. Using this information, we can better differentiate between posts used to recall traumas compared to those used to describe trauma in the abstract.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
medication	ptsd	hear	triggers	against	trauma
meds	symptoms	sound	trigger	beat	happened
medications	depression	sounds	experienced	in front of	past
mg	diagnosis	hearing	events	himself	remember
prazosin	diagnosed	voice	serious	gun	memories

TABLE XIV. INFORMAL TOPICS FOUND IN ASSOCIATION RULES.

Association rules can also be used for identifying and resolving ambiguity in word usage. Consider clusters 7 and 8 found in Table XV. Cluster 7 describes sexual abuse of a child. Cluster 8 describes violent physical abuse. Both of these have association rule connecting them to Cluster 6 which describes "memories" and the "past". The rule $7 \implies 6$ has a lift of 2.42, meaning that users discussing childhood sexual abuse are likely to speak about traumatic memories. The rule $8 \implies 6$ has a lift of 3.02, which implies that users expressing violent abuse of a child are more likely to express past trauma. The rule $5 \implies 6$ has an even higher lift, 3.59. It seems that people describing physically violence are even more likely to post discussions about their memories. These rules taken together show that the term "trauma" in cluster 6 has several interpretations which can be resolved by exploring which topics co-occur with it within a post.

Cluster 7	Cluster 8
abuse	child
childhood	physically
rape	emotionally
physical	violent
sexual	abusive

TABLE XV. INFORMAL TOPICS RELATED TO TRAUMA.

In the previous example, we looked at how several topics might predict a single topic. In this example, we look at the case where a single example predicts several other topics. See Table XVI. Cluster 9 described the physical manifestation of stress. This topic is quite predictive of other topics. The rule $9 \implies 10$ has a confidence of 0.81 and lift of 5.88, meaning that people writing about their physical reaction to stress often talk about their mental reaction to stress. The rule $9 \implies 11$ has a confidence of 0.75 and lift of 5.49, meaning that these same authors often discuss the emotional state such as "angry", "scared" or "tired". The rule $9 \implies 12$ has a confidence of 0.83 and a lift of 6.07, meaning that these authors will often speak about how their stress makes "everyday" a "struggle".

Evaluating the association rules in this manner enables the exploration of comorbidity among the symptoms expressed in PTSD forums. The physical expression of stress seems to be tightly coupled with other symptoms of PTSD such as mental anguish, anger and the feeling of hopelessness.

Cluster 9	Cluster 10	Cluster 11	Cluster 12
body	mind	feeling	hard
heart	lost	scared	avoid
chest	completely	angry	everyday
stomach	my own	tired	struggle
breathing	brain	makes me	escape

TABLE XVI. INFORMAL TOPICS RELATED TO STRESS.

In this section, we demonstrated that by using association rule mining we were able to determine the semantic relationships between the topics. Furthermore, we were able to use these association rules to compare similar seeming clusters to each other, and combine these association rules with clusters in a way which constructs meaningful narrative illustrating common themes in the r/PTSD subreddit. Ultimately, these results indicate that association rules can serve as a useful tool for extracting addition meaning out of existing topics.

F. Discussion

In our assessment of our language model, we first evaluated the automatically generated latent topics. Next, our topics were compared to the criteria used by experts to classify PTSD. Finally, we use an association rule analysis to further explore and validate our clusters. Our analysis revealed several key findings.

First, topic modeling was successful at identifying meaningful topics in the subreddit r/PTSD. This finding suggests the Word2Vec may be generally applicable to other mental health issues and invites the comparison to other topic modeling techniques.

Second, when comparing our topics to the criteria included in the DSM, we found that in some cases we were able to automatically reproduce some of the criteria. In other cases, it was more challenging, particularly when the criteria involved symptoms and experiences over time or symptoms the individual might not express explicitly.

Third, the informal topics we discovered captured many ideas not mentioned in the DSM. Users expressing their struggle with PTSD often discuss the impact it has on their daily lives. Thus we witnessed topics capturing the notions of school, work, travel, weather, and family. The occurrence of these topics in the collaborative narrative of thousands of user's experience with PTSD underscores the pervasiveness of PTSD in the individuals' lives.

Finally, we expanded on previous research by adding analysis using association rules. These association rules proved to be capable of extracting useful information about the context of a word cluster, such as the ways in which the word cluster is used in the corpus. These rules also allow us to compare different clusters, to see which words are most likely to be used to talk about traumatic experiences.

The contribution of the paper is the discovery and analysis of informal latent topics within textual data which is known to contain PTSD. Previous analysis of social media used regular expressions, and general language analysis tools to analyze word use. Our method uses topic modeling to uncover informal latent topics directly from social media posts, which captures the topics and words used by the online community specifically affected by PTSD. Additionally, we analyze these topics to discover the context in which they are used. In the future, these models can be leveraged to monitor the occurrence of PTSD and evaluate changes in the way people talk about PTSD over time.

VI. CONCLUSION

In this work, we trained word embeddings on online forums concerning post traumatic stress disorder. We then clustered words to create informal latent topics. We subjectively evaluated the topics and compared them to criteria described in the Diagnostic and Statistical Manual of Mental Disorders. Finally, we analyzed the relationships between topics by leveraging association rule mining.

Our future work aims to examine other mental health concerns such as depression, eating disorders, and attention-deficit/hyperactivity disorder. Using the models from across different disorders, we plan to evaluate their comorbidity. Finally, we plan to compare and contrast other topic modeling approaches.

ACKNOWLEDGMENT

This work was possible in part by the National Science Foundation under the Research Experience for Undergraduates grant number 1659836. We thank Jason Baumgartner of www.pushshift.io for his archives of Reddit, which helped enable this work. We also would like to express our gratitude to the moderators of r/PTSD for their invaluable service to the r/PTSD community.

REFERENCES

[1] R. C. Agarwal, C. C. Aggarwal, and V. Prasad. Depth first generation of long patterns. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 108–118. ACM, 2000.

[2] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A. I. Verkamo, et al. Fast discovery of association rules. *Advances in knowledge discovery and data mining*, 12(1):307–328, 1996.

[3] J. Alvarez-Conrad, L. A. Zoellner, and E. B. Foa. Linguistic predictors of trauma pathology and physical health. *Applied Cognitive Psychology*, 15(7), 2001.

[4] S. S. Anand, A. Patrick, J. G. Hughes, and D. A. Bell. A data mining methodology for cross-sales. *Knowledge-Based Systems*, 10(7):449–461, 1998.

[5] A. P. Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[6] P. Bamidis, G. Coppersmith, B. Spitzberg, L. LItman, M.-H. Tsou, S. Konstantinidis, S. R. Braithwaite, C. Giraud-Carrier, J. West, M. D. Barnes, and C. L. Hanson. Validating machine learning algorithms for twitter data against established measures of suicidality. *Journal of Medical Internet Research Mental Health*, 3(2), 2016.

[7] J. I. Bisson. Post-traumatic stress disorder. *Occupational medicine*, 57(6):399–403, 2007.

[8] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, Mar. 2003.

[9] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[10] G. Coppersmith, M. Dredze, and C. Harman. Quantifying mental health signals in twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60, Baltimore, Maryland, USA, June 2014. Association for Computational Linguistics.

[11] G. Coppersmith, R. Leary, E. Whyne, and T. Wood. Quantifying suicidal ideation via language usage on social media. In *Joint Statistics Meetings Proceedings, Statistical Computing Section*. JSM, 2015.

[12] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.

[13] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz. Predicting depression via social media. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. AAAI, July 2013.

[14] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2098–2110, New York, NY, USA, 2016. ACM.

[15] S. R. D. C. T. S. Doddi, Achla Marathe. Discovery of association rules in medical data. *Medical informatics and the Internet in medicine*, 26(1):25–33, 2001.

[16] M. Duggan and A. Smith. 6% of online adults are reddit users. *Pew Internet & American Life Project*, 3:1–10, 2013.

[17] G. Gkotsis, A. Oellrich, T. Hubbard, R. Dobson, M. Liakata, S. Velupillai, and R. Dutta. The language of mental health problems in social media. In *The Third Computational Linguistics and Clinical Psychology Workshop (CLPsych)*, pages 63–73, 6 2016.

[18] A. J. Glass and F. D. Jones. Psychiatry in the us army: Lessons for community psychiatry. Technical report, UNIFORMED SERVICES UNIV OF THE HEALTH SCIENCES BETHESDA MD F EDWARD HEBERT SCHOOL OF MEDICINE, 2005.

[19] A. Globerson, G. Chechik, F. Pereira, and N. Tishby. Euclidean embedding of co-occurrence data. *Journal of Machine Learning Research*, 8(Oct):2265–2295, 2007.

[20] D. H. Goh and R. P. Ang. An introduction to association rule mining: An application in counseling and help-seeking behavior of adolescents. *Behavior Research Methods*, 39(2):259–266, May 2007.

[21] J. Han, J. Pei, Y. Yin, and R. Mao. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data mining and knowledge discovery*, 8(1):53–87, 2004.

[22] D. J. Hand and G. Blunt. Prospecting for gems in credit card data. *IMA Journal of management Mathematics*, 12(2):173–200, 2001.

[23] G. A. C. C. T. Harman and M. H. Dredze. Measuring post traumatic stress disorder in twitter. *In ICWSM*, 2014.

[24] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[25] T. Hofmann. Probabilistic latent semantic analysis. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 289–296. Morgan Kaufmann Publishers Inc., 1999.

[26] L. Hong and B. D. Davison. Empirical study of topic modeling in twitter. In *Proceedings of the first workshop on social media analytics*, pages 80–88. ACM, 2010.

[27] S. Ivkovic, J. Yearwood, and A. Stranieri. Discovering interesting association rules from legal databases. *Information & Communications Technology Law*, 11(1):35–47, 2002.

[28] J. Jashinsky, S. H. Burton, C. Hanson, and T. Argyle. Tracking suicide risk factors through twitter in the US. *Crisis The Journal of Crisis Intervention and Suicide Prevention*, 35(1):1–9, 2013.

[29] M. Kumar, M. Dredze, G. Coppersmith, and M. De Choudhury. Detecting changes in suicide content manifested in social media following celebrity suicides. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, pages 85–94. ACM, 2015.

[30] R. Lebret and R. Collobert. Word emdeddings through hellinger pca. *arXiv preprint arXiv:1312.5542*, 2013.

[31] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[32] O. Levy and Y. Goldberg. Neural word embedding as implicit matrix factorization. In *Advances in neural information processing systems*, pages 2177–2185, 2014.

[33] O. Levy, Y. Goldberg, and I. Ramat-Gan. Linguistic regularities in sparse and explicit word representations. In *CoNLL*, pages 171–180, 2014.

[34] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[35] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

[36] B. O'Dea, S. Wan, P. Batterham, A. Calear, C. Paris, and H. Christensen. Detecting suicidality on twitter. *Internet Interventions*, 2(2):183–188, 2015.

[37] R. O'Kearney and K. Perrott. Trauma narratives in posttraumatic stress disorder: A review. *Journal of traumatic stress*, 19(1):81–93, 2006.

[38] C. H. Papadimitriou, H. Tamaki, P. Raghavan, and S. Vempala. Latent semantic indexing: A probabilistic analysis. In *Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, pages 159–168. ACM, 1998.

[39] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[40] R. Řehůřek and P. Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA. http://is.muni.cz/publication/884893/en.

[41] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Commun. ACM*, 18(11):613–620, Nov. 1975.

[42] R. Socher, J. Bauer, C. D. Manning, and A. Y. Ng. Parsing with compositional vector grammars. In *ACL (1)*, pages 455–465, 2013.

[43] R. Socher, A. Perelygin, J. Y. Wu, J. Chuang, C. D. Manning, A. Y. Ng, C. Potts, et al. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, volume 1631, page 1642, 2013.

[44] C. Wang and D. M. Blei. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 448–456, New York, NY, USA, 2011. ACM.

[45] L. Wolf, Y. Hanani, K. Bar, and N. Dershowitz. Joint word2vec networks for bilingual semantic representations. *Int. J. Comput. Linguistics Appl.*, 5(1):27–42, 2014.

[46] G. Zanotti, M. Horvath, L. N. Barbosa, V. T. K. G. Immedisetty, and J. Gemmell. Infusing collaborative recommenders with distributed representations. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pages 35–42. ACM, 2016.