"I Will Not Drink With You Today": A Topic-Guided Thematic Analysis of Addiction Recovery on Reddit

Robert P. Gauthier
University of Waterloo
Waterloo, Ontario, Canada
rpgauthier@uwaterloo.ca

Mary Jean Costello
Homewood Research Institute
Guelph, Ontario, Canada
JCostello@homewoodhealth.com

James R. Wallace
University of Waterloo
Waterloo, Ontario, Canada
james.wallace@uwaterloo.ca

ABSTRACT

Recovery from addiction is a journey that requires a lifetime of support from a strong network of peers. Many people seek out this support through online communities, like those on Reddit. However, as these communities developed outside of existing aid groups and medical practice, it is unclear how they enable recovery. Their scale also limits researchers' ability to engage through traditional qualitative research methods. To study these groups, we performed a topic-guided thematic analysis that used machine-generated topic models to purposively sample from two recovery subreddits: *r/stopdrinking* and *r/OpiatesRecovery*. We show that these communities provide access to an experienced and accessible support group whose discussions include consequences, reflections, and celebrations, but that also play a distinct metacommunicative role in supporting formal treatment. We discuss how these communities can act as knowledge sources to improve in-person recovery support and medical practice, and how computational techniques can enable HCI researchers to study communities at scale.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; **User studies**; • **Computing methodologies** → **Machine learning**.

KEYWORDS

Reddit; machine learning; thematic analysis; addiction; online communities

1 INTRODUCTION

Recovery from substance addiction [1] can involve long and difficult journeys [99]. A key component of those journeys is having a strong network of peers who can support a person as they work

towards a healthy, productive, and meaningful life [11, 17, 48, 53]. Common sources of this support are health professionals, rehabilitation programs, and 12-step programs like Alcoholics Anonymous (AA) or Narcotics Anonymous (NA). However, several barriers can prevent people from participating in these groups, like physical distance, lack of cultural similarity to peers, the stigma surrounding addiction, and a program's appeal [28, 32, 48, 64]. As a result, many seek out less formal support communities through social networking platforms like Reddit or Facebook [39, 67]. However, since these communities have developed outside of existing support groups and clinical practice, it is unclear whether they provide appropriate and effective support to those who seek it.

Our work investigates what is discussed in online peer-to-peer communities that have formed around addiction recovery and contributes an understanding of how these communities support their members' recovery journeys. We focused our inquiry on the use of Reddit (www.reddit.com), a pseudonymous social networking site where communities can discuss sensitive topics that people may not feel comfortable disclosing face to face or on sites were they are personally identifiable, like Facebook [63]. We investigated discussions in two subreddits, r/stopdrinking and r/Opiates-Recovery, where people can seek out advice about recovery from addiction to two common substances, alcohol and opiates [22].

To understand how these online communities support recovery, we built on previous Human-Computer Interaction (HCI) research (e.g., [2, 3, 82, 83]) by developing computationally-supported qualitative research methods [71]. In particular, we addressed key drawbacks of existing research in this space, arising from the tension between the time intensity of qualitative research goals and the scale of online communities. That is, existing research has commonly resorted to: 1) sampling only a small set of posts from each community to enable human researchers to develop a qualitative understanding of the materials (e.g., [98]); or 2) focusing on quantitative analysis (e.g., [2]) and losing much of the 'thick' understanding of these communities [18].

In this work, we applied computational techniques to perform a 'topic-guided thematic analysis' of discourses on recovery subreddits. First, we used Latent Dirichlet Allocation (LDA) [16], an unsupervised topic modelling technique, to develop models for each subreddit from four years of posts. We used these models' topics to generate purposive samples [29, 42] by identifying related keywords and representative threads from both subreddits. We then performed reflexive thematic analysis [18] on our purposive samples' threads to develop and review our themes. During our analysis, we performed inductive coding and grounded our interpretations in the communities' original contexts by looking at associated threads

on Reddit's website. This combination of unsupervised topic modelling, purposive sampling, and reflexive thematic analysis enabled us to develop qualitative understandings of these communities while sampling from more than 150,000 threads. We present the results of our analysis in terms of two research questions: 1) How are stories used for addiction recovery in these Reddit communities?; and, 2) How do community members support each other's recovery?

Our research contributes an empirical understanding of the discussions people on recovery journeys are having online, and the information they have sought and shared on social networks at a large scale. We show that the communities comprise experienced members, are perceived as accessible, and provide a channel for sharing lived experiences such as personal stories, advice on common problems, and emotional support. These resources are leveraged by people experiencing addictions, their family, and their friends. Further, we show that Reddit enables meta-discussions that help people from under-represented groups navigate in-person programs, for example women seeking women mentors, and people struggling with references to 'a higher power' in 12-step programs. In discussing these findings, we describe how our themes provide a holistic understanding of addiction recovery that includes online communities, and can inform practice for both mutual aid programs and healthcare practitioners. Finally, we reflect on the effectiveness of computational techniques in supporting the development of a qualitative understanding of online communities.

2 RELATED WORK

Experts today view addiction recovery as an ongoing journey that requires a variety of supports to enable those in recovery to "...develop a healthy, productive, and meaningful life." [99, p. 236]. While professional treatment programs vary in implementation, they often share common elements such as an emphasis on education, development of coping skills, and management of co-occurring symptoms such as post-traumatic stress disorder [1]. In many programs, emphasis is placed on support and mentorship through mutual aid groups, which may involve participation in well known 12-step groups, like AA and NA, or alternative groups, such as Self-Management And Recovery Training (SMART) and Moderation Management (MM) [28, 48]. That is, the formation of a lasting, positive, behaviour-dependent, and supportive network of peers is considered a key component of long-term success [11, 17, 48, 53].

There is a growing body of work in the HCI literature that examines how structured online health communities can provide this support network, facilitate emotional support and information exchange [36], and help their members manage health challenges (e.g., [47, 60]). Notably, these structured communities have been found to be helpful for those with chronic conditions, including addiction [12, 80, 84, 92, 100, 105, 106]. To date, this research has largely focused on structured communities, such as InTheRooms [12, 86, 87] and MedHelp [25, 58], and has found that they offer similar benefits to in-person support groups, such as AA and NA [12].

However, these structured online health communities also have drawbacks. They are often accessible only to those with registered accounts, potentially deterring people, and their friends and family,

from accessing needed support. They also may place an emphasis on certain topics, perspectives for successful treatment, or belief systems that do not work for everyone [48, 54]. For example, InThe-Rooms places an emphasis on the 12-step programs AA and NA [49], whereas MedHelp's focus on connecting people with medical professionals emphasizes experts' opinions [56]. For these reasons, there is growing interest in understanding what kind of support open-access online health communities on platforms like Reddit provide their members.

Further, the scale of these open-access online communities makes them difficult to study. Previous HCI research has begun exploring approaches to computer-supported qualitative research methods (e.g., [2, 3, 24, 35, 71]). Nevertheless, researchers investigating addiction tend to make one of two compromises in their approach: 1) they sample a small subset of posts from each community to enable human researchers to develop a qualitative understanding of the materials (e.g., [33, 93, 98]), limiting their studies from including the years of discussions that members of the communities have access to, or 2) they focus on quantitative analysis (e.g., [82, 83, 95]) limiting much of the 'thick' understanding of these communities that qualitative analysis could have developed [18].

In our work, we make two contributions towards understanding the benefits of online communities focused on recovery: We first demonstrate use of computational methods for qualitative analyses of online communities' discussions, called a 'topic-guided thematic analysis', to overcome limitations of existing methods. We then perform an analysis on two addiction recovery communities on Reddit, describe how they support addiction recovery, and show how they help people from under-represented groups navigate in-person programs.

2.1 Analysis of Online Communities, Computational Support, and Sampling

One of the most significant challenges of studying online communities is their scale: each community is potentially comprised of hundreds of thousands of posts from tens of thousands of people over a period of years. Thematic analysis is a time-intensive method, due the the amount of reading, re-reading, and reviewing involved, that aims to explore and develop understandings of complex data [18]. The scale of large online communities both amplifies the amount of time needed for analysis and makes finding data that contain interesting aspects difficult. Research to date often overcomes this challenge through different approaches to *sampling*; such as, selecting posts from a small time period [98] or those associated with 'hot topics' at a given point in time [33]. For instance, Wadley et al. [98], in their investigation of r/StopSmoking chose a sample frame of the 732 posts made during April 2014, and then randomly selected 100 posts from within that sample for manual coding. While these sampling approaches enable manual coding of the data, they also have substantial limitations: for instance, they limit researchers' opportunity to familiarize themselves with the data and develop a contextual understanding of the communities, and can exclude data from dominant or seasonal trends [62].

Computational techniques provide an opportunity to overcome these sampling limitations, and enable an in-depth, qualitative understanding of online discourse [35, 71]. For example, Latent Dirichlet Allocation (LDA) [16], an unsupervised modelling approach that can identify latent topics and associated threads within an online community [59], can be used to purposively sample [29, 42] discourses for analysis. Researchers have employed the use of computational methods to derive a variety of topics from large data corpi for some time (e.g. [8, 30, 31, 34, 68, 74]). LDA has been particularly useful to HCI researchers in identifying latent topics in Reddit communities (e.g., [2, 3, 82, 83]), and has been described by Ammari et al. [3] as part of a 'roadmap' for using computational techniques to better understand social relationships online.

Inspired by their work, we further develop this roadmap with a focus on a qualitative, *human* understanding of online discourse. We applied computational techniques to perform a 'topic-guided thematic analysis' of discourses on Reddit, where unsupervised LDA is used to sample for a thematic analysis. Our approach parallels explanatory mixed methods designs [29], and explores the use of computational techniques to augment human researchers' abilities, as described by Muller et al. [71].

The use of LDA to purposively sample for thematic analysis has several advantages. First, our two LDA models use the breadth of each corpus of threads ($144, 422$ threads from r/stopdrinking and $14, 079$ threads from r/OpiatesRecovery) which allows topics to emerge from all threads, rather than from a small sample. Second, LDA assumes that each thread comprises a mixture of topics, enabling the identification of secondary and/or latent topics [16], which aligns with how Reddit threads can involve multiple people contributing different viewpoints on both initial posts and subsequent responses. Third, purposive sampling via LDA enables identification of multiple threads for each topic, providing us opportunities to iteratively validate the models for human semantic sense [59] and to identify the samples needed for thematic analysis.

2.2 Open-Access Online Communities and Addiction Recovery

Large communities exist on open-access platforms around issues like addiction [33]. In particular, research has shown a large degree of participation on social networking platforms for topics like smoking cessation [98] and diabetes [76]. Previous work has also shown that online communities possess the same treatment mediators present for in-person group support [33]. For instance, Q&A participation in online communities has been found to be motivated by altruism and efficacy; known mediators for in-person mutual aid groups that are associated with an increased likelihood of recovery [79]. As such, these open-access online communities can be considered to be mutual aid groups.

Additionally, pseudonymity fosters disclosure in online forums, particularly for sensitive topics like addiction [89, 104], diabetes [76], and pregnancy loss [4]. That is, people are often less willing to discuss sensitive topics when those discussions are linked to their real identities on platforms such as Facebook or LinkedIn [76]. On the other hand, platforms that support or even encourage various degrees of anonymous participation, like Reddit, have been found to be supportive of these sensitive discussions [3, 5, 72]. The

freedom to post anonymously has been shown to enable discussions of mental health [6, 14], parenting issues [3], and sexual abuse [5].

However, the peer-to-peer nature of these open-access online mutual aid groups, and their lack of medical authority, raises questions surrounding whether the information and advice are appropriate in the context of addiction [33, 79]. That is, should these forums be officially recommended by health professionals as a place to find support? Initial work has suggested several limitations of online communities. For instance, community members may largely be new to sobriety, and advice provided by these groups may be harmful rather than helpful [27, 93]. Similar concerns arise from the perspective of members seeking help, for instance Rubya and Yarosh [87] found that community members may be less likely to self-disclose information in online meetings than when meeting face-to-face, and that geographic differences make communication less effective, or may even lead to conflict. Barrett and Murphy [7] found that online meetings were not perceived as being more accessible than face-to-face meetings, and were perceived as being less effective and of lower quality.

Further, much of the work that seeks to understand these communities is focused on the perspective of clinicians, such as looking for themes from diagnostic tools [33, 37], clinician expertise-based recommendations [47], analyzing discussions to find addiction mechanisms and treatment methods that can be assessed for clinical validity [23, 51], and categorizing users into clinical diagnoses [57, 95, 107]. While valuable, clinician perspectives may not always align with community members' values [13].

To address this gap in our understanding, we performed analyses of two active addiction-support subreddits. We show that the topics discussed align well the support identified in the healthcare literature [11, 17, 48, 53], and are consistent with a positive, supportive network of peers. Further, we found that these open-access online communities serve a meta-communicative role in helping people to navigate difficulties with in-person groups, such as women seeking women mentors, and people struggling to accept religious aspects of AA.

3 METHOD

We chose to study two subreddits because we initially sought to establish both similarities and differences across recovery communities. However, as our analysis progressed, it evolved to focus on exploring themes that were generated from our interpretation of both subreddits' discussions. Such an evolution is expected when performing thematic analysis using an inductive coding process [19].

We investigated two active addiction recovery subreddits focused on alcohol addiction (r/stopdrinking) and opiate addiction (r/OpiatesRecovery), retrieving corpa from pushshift.io [10]. These subreddits were selected because: they address recovery from use of two different classes of substances that are stereotyped as legal (alcohol) and illegal (opiates) and are current concerns of public health [81], they were the largest recovery subreddits we could find for each substance, they are publicly accessible, and they are active in terms of number of community members and posts.

We used thematic analysis to create in-depth understandings of the behaviour we observed in both subreddits. Thematic analysis enabled us to develop 'thick' understandings of the subreddits'

community discussions and to generate results that are accessible for both researchers and the general public, while also capturing unanticipated insights [18]. To perform the thematic analysis, we first built LDA models for each subreddit and used those models to purposively sample [29, 42] threads for each topic. Our LDA models used the full corpus of texts (144, 422 threads from r/stopdrinking and 14, 079 threads from r/OpiatesRecovery) to develop 16 topics for each subreddit. 20 discussion threads were selected from each topic in each of our two 16-topic LDA models, providing a total sample of 640 threads (composed of 640 submissions and 7828 comments). We then performed a reflexive thematic analysis [18] where we inductively coded the sampled discussions, in their living state on Reddit, and used our codes and samples to develop our themes.

Our iterative approach comprised three phases: data gathering, LDA topic modelling, and thematic analysis (Figure 1). The code used to support each activity was written in Python, and is included in the supplementary material.

3.1 Data Gathering and Ethical Considerations

As addiction recovery is a sensitive topic, we took additional steps to consider the ethical implications of our work and to protect the communities that we were interested in learning from. We reviewed Reddit's terms of service and the rules and FAQs of both of the subreddits to confirm that data was open for public use, though they did not explicitly allow use for research purposes. We included all publicly available threads from the r/stopdrinking and r/OpiatesRecovery subreddits created during 2014 to 2017 (144, 422 threads and 14, 079 threads respectively). We chose the start point of January 1, 2014 and the end point of December 31, 2017 based on our desire to allow behaviours that might be seasonal to occur multiple times and what was available when we started performing the analysis in the summer of 2018. Data for submissions and comments was downloaded from pushshift.io in json format [10].

We used discussion threads as our unit of analysis because we intended to examine entire threads for our thematic analysis, to preserve the context of each submission and its responses. Threads were recreated by merging title and text fields by submission id for both the submission and any associated comments. The downloaded data was also used to identify URLs as well as aggregated to find thread counts and distinct user id counts by date. Other non-aggregate information was not extracted or used from the dataset to respect both the community members' privacy (i.e., classifying and categorizing community members risks inferring private information [97]) and limitations of pushshift.io [10] as a snapshot of posts on Reddit (e.g., archived karma scores often do not match those currently on the website).

All published quotes are paraphrased from existing non-deleted posts to preserve pseudonymity. To respect the choice of community members who chose to delete content, we did not include deleted posts in our thematic analysis. Deleted content was identified manually when reviewing threads on Reddit for thematic analysis. We paraphrased by breaking quotes down into their thematic analysis codes, then manually constructed a new quote. We

then compared against the old quote for consistency, and Googled it to ensure anonymity.

Our study design received full approval from our institution's research ethics board, and is consistent with guidelines from the HCI community for protecting pseudonymous research participants (e.g., [20, 61]) and transparency in qualitative research [94].

3.2 LDA Topic Modelling

LDA topic modelling involved three iterative sub-phases: cleaning, building, and evaluation. These sub-phases ultimately produced two models: one for each subreddit. Each model comprised 16 topics, keywords, and a list of associated threads that we used as purposive samples for our thematic analysis. Although we cannot release the final cleaned datasets or generated models, since they contain non-paraphrased text data that could be used to identify community members, our code and a summary of outputs is available in Appendices A and B.

3.2.1 Cleaning. Our cleaning process emphasized human interpretablity [9, 59], an important consideration given our goal was to create purposive samples for thematic analysis by a human researcher (e.g., [9, 71]). Our initial cleaning approach consisted of: 1) lemmatization and part of speech identification [50] using spaCy [44]; and 2) English stop word filtering, as well as filtering out words that were not nouns, verbs, adjectives, and adverbs [50] using NLTK [15] and Mallet [65]. This kind of 'light cleaning' improves the interpretability of models, but does not impact their stability [90, 91].

In subsequent cleaning iterations we noticed repetition of common acronyms within models' topics. To reduce this duplication and allow the LDA modelling to treat all representations as being the same word, we expanded a number of general and domain-specific acronyms (e.g., 'fyi' to 'for your information' and 'wd' to 'withdrawal') to their full form. We also observed that frequently occurring words, particularly adjectives, were causing a high level of overlap between topics. To reduce this overlap, we removed both adjectives and words that appeared in more than 25% of threads. To further improve performance, we also removed words that occurred in fewer than 20 threads using Gensim [85] and masked out external links (e.g., 'http://....').

3.2.2 Building. We built an LDA model for each subreddit using Gensim 3.8 [85]. For metadata parameters we set the number of passes to 100 along with both alpha and eta to auto, to allow each model to infer its own asymmetric topic distribution from the corpora [85]. We then built 10 separate LDA models for each subreddit, using the generated dictionaries and corpora, and selected the model with the maximum coherence score from the 10 generated models (Appendix A and B).

We set the number of topics in each model to 16, based on pilot runs that indicated the topic coherence had plateaued; previous research has found that topic models with higher coherence score correlate to human-interpretable topic groups [59, 75, 88]. We selected 16 topics for each subreddit because, although opiates recovery plateaued earlier at 9 topics, we wanted to gather a similar sized sample from each subreddit to allow our thematic analysis to consider each community equitably (Appendix C).
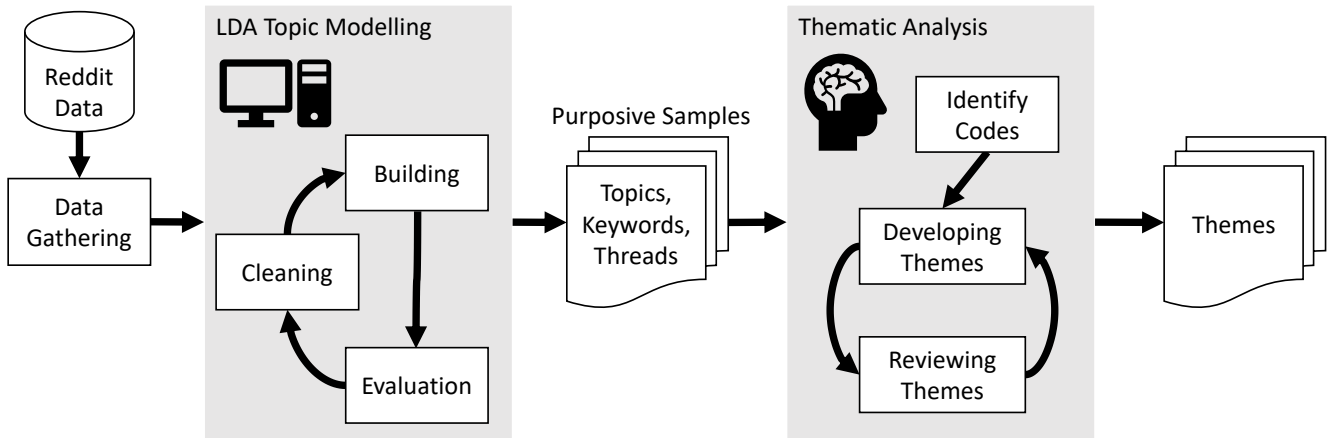
Figure 1: To perform our thematic analysis, we first collected Reddit data from 2014 to 2017 for `r/stopdrinking` and r/Opiates-Recovery. We then developed LDA topic models for each subreddit by iteratively cleaning, building, and evaluating the generated models, which produced proposive samples made up of keywords and threads for each topic. Finally, we performed the core of our thematic analysis by iteratively identifying codes, developing themes, and reviewing themes with the topics, keywords, and threads from our LDA model.

3.2.3 Evaluation. We then iteratively fine-tuned the LDA topic models [59]. We used the models to categorize the available threads and inspected the distribution between different topic groups. We reviewed the models, their topic terms, and the topics' associated threads to assess their reliability [59], coherence (i.e., C_V [88]), and, most importantly, whether the topics were interpretively useful [9] to both the authors and a separate pilot group of 12 HCI researchers. During these reviews we reflected on whether additional cleaning was required. During early iterations, we also tried using Jaccard's distance as a measure of topic similarity [3]. However, as we adjusted our cleaning process the Jaccard's distance measures approached 1.0 for almost all identified topics, and so we did not ultimately use it to guide our topic modelling. Instead, we relied on our semantic interpretation of generated topics, consistent with our need to inform our thematic analysis.

3.2.4 Purposive Sampling. To purposively sample we used each subreddit's LDA model to retrieve 20 representative `submission ids` for each of the 16 topics, giving a total of 640 threads for analysis. We identified representative threads by calculating the probability of each topic occurring in each thread, and selected the 20 threads with the highest probability. We then generated URLs for each `submission id` that could access the discussion threads. We accessed each thread through Google Chrome to ground our analysis in the context of the Reddit communities. The 320 `r/stopgaming` threads were composed of 3302 comments with a median of 7 comments per thread (mean = 10.32). The 320 `r/OpiatesRecovery` threads were composed of 2526 comments with a median of 6 comments per thread (mean = 7.89).

3.3 Thematic Analysis

For our reflexive thematic analysis [18] we took the realist stance that the continued existence of these subreddits implies a perception of value by their communities and that seeking to understand

addiction recovery is a complex process with many different contributing processes and possible outcomes. To focus on the experiential knowledge of the subreddits, we used inductive coding rather then try to force the behaviour of the communities into current understandings of addiction recovery or categorize the communities' members.

We based our reflexive thematic analysis approach on the 6 phases described by Braun and Clarke [18]. The familiarization phase occurred as part of the LDA topic modelling. Using what was learned from familiarization, the first author then worked with the purposive samples to identify codes and to develop and review themes. The first author read each thread in its original context on Reddit to understand what data-driven codes were present. The first author iteratively compared the threads and codes to develop themes (Appendix D); for example, the common code pair 'seeking information' and 'providing information' was combined with codes for different types of seekers, such as 'Atheists' and 'Female', and support group sub-codes, such as 'overwhelmed' and 'fear of stigmas' to develop the theme 'Navigating 12-Step programs'. Threads could contain multiple codes, codes could contribute to multiple themes, and thus threads could also contribute to multiple themes. The first author gathered supporting quotations for each theme from multiple threads that contributed supporting codes. Finally, the first and third authors reviewed the themes by looking at additional threads from each subreddit to confirm that they were present and came to agreement that the themes fit the data.

To consider the individual researcher positions inherent in qualitative research, such as our reflexive thematic analysis, we conducted group reflections with the first, second, and third authors on the identified themes [29]. During these group reflections we discussed the topics used for sampling, the themes developed, the example quotes, and the first author's experiences and thoughts from interacting with content on a sensitive subject. We did not