

Exploring the Landscape of Drug Communities on Reddit: A Network Study

Layla Bouzoubaa

College of Computing & Informatics
Drexel University
Philadelphia, USA
lb3338@drexel.edu

Jordyn Young

College of Computing & Informatics
Drexel University
Philadelphia, USA
jby28@drexel.edu

Rezvaneh Rezapour

College of Computing & Informatics
Drexel University
Philadelphia, USA
sr3563@drexel.edu

Abstract—In a world where the use of social media is almost universal, understanding its influence on drug information consumption has become a crucial topic of research. This study delves into the networks of drug-centric subreddits to gain insights into community structures and the nature of user interactions within Reddit. After conducting a thematic analysis of 131 drug-related subreddits, we selected four dominant ones for a comprehensive network analysis, examining the broader interactions of their active users. Our results show that drug-related subreddits are highly connected to each other, resembling small-world networks, and tangential subreddits, like *r/AskReddit* and *r/gaming*, are influential in connecting different parts of networks. While each community within the sub-networks has its unique focus, there is a clear convergence around shared themes and interests, notably harm reduction and mutual support. Our research not only underscores the pivotal role of online platforms like Reddit in bridging the active members of drug-related subreddits to a broader spectrum of content but also highlights the underlying commonalities that weave these communities together, emphasizing their mutual interests and shared values.

Index Terms—social network analysis, Reddit, substance use, topic modeling, community analysis

I. INTRODUCTION

The world lost 1.25 million years of life to unintentional drug overdose deaths (ODDs) between 2015-2019 [1]. With the rise of the internet and social media platforms, drug-related information is more available and accessible than ever before, which has implications for both prevention and treatment efforts. To effectively combat ODDs, it is critical to have a comprehensive understanding of both the motives behind individual drug use and the broader context in which users navigate online drug communities. By analyzing these networks of information exchange, we can identify key nodes of influence and dissemination patterns, which can be invaluable for devising targeted support and treatment strategies [2].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

© 2023 Association for Computing Machinery.

ACM ISBN 979-8-4007-0409-3/23/11...\$15.00

<https://doi.org/10.1145/3625007.3629125>

One platform that has gained popularity for drug-related discussions is Reddit. Reddit is a social news aggregation website that allows users to create and participate in communities, called subreddits, that focus on specific topics. Drug-focused subreddits have become popular destinations for individuals seeking information and support related to drug use, dependencies, and recovery [3]. Previous research shows that these online communities can provide a supportive environment for individuals to share their experiences around stigma [4], seek advice on treatment strategies [5], and receive emotional support [6]. Despite the widespread popularity of subreddits that revolve around drugs, little is known about their network structures and community foci. To fill this gap, we performed a comprehensive analysis of drug-related subreddits to categorize their objectives, characteristics, and structures. By analyzing these aspects, we aim to gain a better understanding of the dynamics and interconnections that exist within and between these online communities. For this purpose, we use a combination of qualitative and quantitative methods to answer the following research questions:

- RQ1: What are the characteristics of the network of Reddit users who engage in drug-related discourse?
- RQ2: How do existing communities within this Reddit network connect and interact with each other?

To study overall drug-related discourse on Reddit (RQ1), we used the Reddit API to collect data about 131 drug-related subreddits. After conducting a rigorous thematic analysis of the various subreddits, we identified six distinct categories of subreddits that play a crucial role in shaping discussions related to drug use and its associated risks. These categories are: Drug Information, Harm Reduction, Continued Drug Use, Stop/Reduce Drug Use, Experience Seeking, and Miscellaneous. Among these categories, Continued Drug Use was the largest area of interest, prompting us to conduct a more in-depth examination of subreddits that focused on this topic.

To gain deeper insight into the connections and interactions between subreddits dedicated to Continued Drug Use (RQ2), we performed a network analysis on a sample of subreddits. These subreddits pertain to four distinct categories of drugs based on their effects on the body: depressants (*r/benzodiazepines*), hallucinogens (*r/LSD*), narcotics

(*r/opiates*), and stimulants (*r/cocaine*). These subreddits were also selected to provide a comprehensive representation of the substances that are the most commonly abused and/or misused [7]. To better understand how these subreddits are connected within and across different drug classes, we calculated centrality measures for each subreddit’s ego network and performed community detection using Louvain’s algorithm [8]. The network analysis of the four primary subreddits revealed dense connections and interactions between them, as well as edges to other related and popular subreddits, such as *r/AskReddit* (general Q&A), *r/Drugs* (general discussion on drug use), *r/weed* (cannabis related discussion), and *r/Showerthoughts* (sharing instant realizations). To investigate the themes and contents within each emergent community, we performed topic modeling on the subreddit descriptions.

Our results provide insights into the patterns of communication and information in drug-related communities. Our findings can inform public health efforts by identifying connections and, therefore, influence information-seeking behavior, to provide support to people who use drugs (PWUD) and reduce the harms associated with information on continued drug use.

II. RELATED WORK

A. The Internet, Social Media, and Drugs

Many of the conversations and educational opportunities that were once facilitated by physical interventions are now happening online [9]. This is likely because online groups have lower barriers to entry while still facilitating social connections [3], [10]. Researchers have confirmed the existence of drug-related communities in digital spaces [11], [12] and noted their potential to both harm and help users [13], while emphasizing that our understanding of such online groups is lacking [14]. The current research on drug discourse on social media has been very specific. For instance, some studies investigated how individuals with opioid use disorder (OUD) communicate about treatment drugs [15], recovery [6], or withdrawal [5]. While these targeted investigations are important, there is a lack of comprehensive research that zooms out to analyze the broader landscape of drug discourse on social media platforms. An early study in 2006 did categorize drug discourse on the internet based on information collected from the Psychonaut Project in 2002 [16]. However, the digital landscape of online communications has drastically changed in the last 20+ years, and conversations about drugs have evolved. As such, it is essential to perform a more current and relevant analysis of drug discourse on social media platforms.

B. Reddit

Reddit, a social media platform with over 50 million daily active users, is a valuable source of user-generated data for research on drug use. It attracts users from a wide range of backgrounds and with diverse experiences, including harm reduction [17] and personal experience perspectives [6]. Reddit users are often more anonymous than users on other social media platforms, which encourages honesty and openness in their conversations. This anonymity is especially important for

studying sensitive topics like substance use, where users may be hesitant to share their experiences in public [18].

Access to data about conversations within subreddits can be used to gain insights into a range of topics, from public opinion [19] to social support trends [20] to health behaviors [21], and inform policy and practice in a variety of fields. Beyond the research community, policymakers also benefit from studying such organically generated online data in order to create and implement effective change [22]. While researchers frequently use Reddit to study drug-related discourse, a limitation of these studies is that they often look at only one drug or one community at a time. In our study, we build on the previous literature by exploring cross/inter-community characteristics.

C. Networks

Network analysis offers insights into relationships within communities [23]. Prior research has shown that the structure and function of online networks are associated with personal drug use. For example, denser networks are associated with higher levels of individual substance use [24]. However, the relationship between online networks and substance use varies by the type of substance used [24]. Youth who exhibit social isolation are particularly more susceptible to increased substance use [25]. Copeland et al. (2018) suggest that youth who are disengaged (not sending ties) and outside-oriented (having close out-of-school ties) are at increased risk for cigarette and marijuana use [25]. Evidence suggests a link between online substance use discussions and personal usage, making it vital to understand the characteristics of online drug communities.

III. METHODS

A. Subreddit Selection

We selected the subreddits for our study by referencing a Reddit post [26] dedicated to compiling a list of subreddits covering various topics, including drugs and drug-related discussions. This list initially comprised 175 subreddits. We then conducted a manual review to identify subreddits that primarily featured text content (versus audio or image-only content), which reduced the list to 141 relevant subreddits. We then utilized the Reddit application programming interface (API)¹ to extract descriptive information from these subreddits. We excluded subreddits that were removed or lacked any content, resulting in a final set of 131 subreddits. More specifically, we utilized the “About” page of each subreddit as the textual data for our analysis. This section of the page provides details on the community’s purpose, rules, guidelines, and moderators. It also includes information about the topics that are relevant to the given subreddit, the types of posts that are allowed or discouraged, and links to related subreddits or external resources that community members may find useful.

B. Thematic Analysis

To identify the intended purpose of each subreddit, we applied a coding scheme introduced by Schifano et al. (2006)

¹[https://oauth.reddit.com/r/\[SUBREDDIT\]/about](https://oauth.reddit.com/r/[SUBREDDIT]/about)

[16] to assess the description of each subreddit. This system uses four main codes:

- Anti-drug: advocating against the misuse of any drug; offering either prevention and/or treatment information
- Pro-drug: actively promoting or facilitating the use of psychoactive drugs; providing information on how to synthesize, purchase or consume a variety of substances
- Harm Reduction: neither condemning nor promoting the use of drugs; presenting evidence and facts, giving the surfer/reader the possibility to draw their own conclusions
- Not Stated: cases in which no statement was found and/or when the subreddit’s aims were unclear.

Using this scheme, two authors annotated the 131 subreddit descriptions. The Cohen’s kappa to measure inter-rater reliability between the two authors was 0.57, indicating moderate agreement [27]. Subsequent conversations between the authors revealed that the agreement was the result of a discrepancy between the codes and the data. The authors agreed that their interpretations of the data did not match the existing coding scheme; rather than promoting or opposing drugs, the community descriptions were human-centered and focused on members instead of a moderator’s opinions. For example, the subreddit *r/benzodiazepines* had the following description: “<https://discord.gg/bonzos>. Come chat with other benzo users on our discord!” The authors recognized that the existing classification system of anti- or pro-drugs, established 20 years ago, ignores the nuanced aspects of this description thus constraining the depth and relevance of any conclusions that can be drawn. Furthermore, the scheme did not include categories for subreddits that discuss the use of specific drugs, such as *r/LSD* or *r/cocaine*. The code for Harm Reduction was also found too broad and did not adequately capture the diversity of subreddits that focus on this topic.

In an effort to improve this coding system, the authors made note of patterns they observed in the data. They concluded that a significant portion of the data could be categorized as either informational or interpersonal. Despite the complexity within these categories, they developed a revised coding scheme consisting of six possible codes:

- Drug Information: scientific discussion around any substance, licit or illicit, that could promote or facilitate research and understanding of various properties of the substance
- Interpersonal Support:
 - Experience Seeking: People with a shared interest for substances coming together to do/discuss other things (e.g., a book club, memes, music) without necessarily consuming drugs
 - Continued Drug Use: Sharing personal experiences and/or information about the substance and promoting community among people who use; “judgment-free”
 - Stop/Reduce Drug Use: Sharing personal experiences/struggles; includes human-centered groups about addiction and recovery

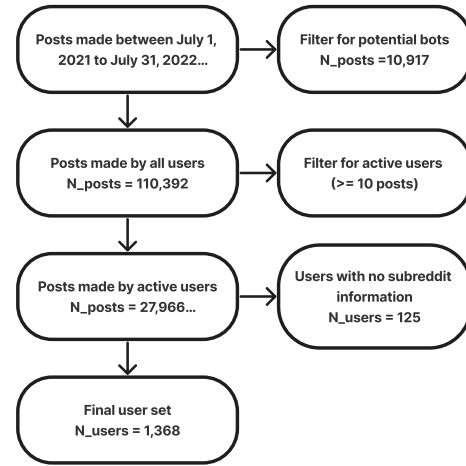


Fig. 1: Criteria for selecting users from whom post activity was gathered for network analysis.

- Harm Reduction: Groups specifically meant to promote safe drug use; engage in discussion around resources and practices for safe drug use
- Miscellaneous: Blank or unrelated

Upon implementing the new coding scheme, there was a significant improvement in the level of agreement between the authors ($\kappa = 0.76$). The new coding scheme enabled the authors to capture more detailed information compared to the previous coding scheme.

C. Network Analysis

To better understand how people discuss and interact on drug-related subreddits, we focused on subreddits that discuss drugs that are commonly abused and misused. We were particularly interested in exploring the links between these subreddits and other subreddits that their active users engage and connect with. Within the Continued Drug Use category, identified through thematic analysis, we selected four subreddits that were the most representative of commonly abused drugs: *r/LSD*, *r/benzodiazepines*, *r/opiates*, and *r/cocaine* [7].

To assess the interactions between these four subreddits and others, we collected information from all active users of these subreddits from July 1, 2021, to July 31, 2022. We defined an active user as someone who had created at least ten posts within this time frame. To ensure that our user sample was free of potential bot accounts, we implemented a two-step process. First, we searched for any variations of the term “bot” in the usernames. We then filtered out any users with negative karma or a post-karma score of zero. Karma, in this context, serves as a measure of post or comment popularity. Figure 1 presents a summary of the final user dataset from which our network edgelist was created. We then identified all the subreddits with which our set of users interacted, beyond our four primary subreddits. This process helped us to establish an extensive network of interactions between the four drug-related subreddits and other subreddits on the platform. In these networks, nodes are subreddits, edges reflect interactions of active users with other subreddits, and edge weights reflect

TABLE I: Summary statistics of the 131 subreddits whose codes were agreed upon by the authors

	Drug Information (N = 15)	Harm Reduction (N = 11)	Continued Drug Use (N = 61)	Stop/Reduce Drug Use (N = 13)	Experience Seeking (N = 17)	Miscellaneous (N = 14)	Total (N = 131)
Number of Subscribers							
Mean (SD)	55,800 (66,600)	50,500 (68,200)	141,000 (256,000)	77,900 (109,000)	77,200 (90,200)	29,100 (76,600)	97,200 (189,000)
Median [Min, Max]	21,100 [908, 203,000]	16,300 [1770, 205,000]	78,700 [369, 1,860,000]	29,800 [1,120, 365,000]	74,500 [463, 321,000]	2,300 [431, 291,000]	43,700 [369, 1,860,000]
Restricted?							
FALSE	12 (80%)	9 (81.8%)	60 (98.4%)	10 (76.9%)	16 (94.1%)	8 (57.1%)	115 (87.8%)
TRUE	3 (20%)	2 (18.2%)	1 (1.6%)	3 (23.1%)	1 (5.9%)	6 (42.9%)	16 (12.2%)
Over 18 Community?							
FALSE	14 (93.3%)	9 (81.8%)	56 (91.8%)	13 (100%)	12 (70.6%)	14 (100%)	118 (90.1%)
TRUE	1 (6.7%)	2 (18.2%)	5 (8.2%)	0 (0%)	5 (29.4%)	0 (0%)	13 (9.9%)

the frequency of interactions between every two subreddits (number of mutual users shared between the subreddits).

After creating the networks of the four primary subreddits: (*r/LSD*, *r/benzodiazepines*, *r/opiates*, and *r/cocaine*), we explored characteristics of their interaction with other subreddits. This involved examining the nodes and edges that directly connect to each subreddit and exploring the patterns of behavior and communication among the subreddits in these networks. We calculated four centrality measures for each network: degree, closeness, betweenness, and eigenvector centrality. For each ego network, nodes with a degree less than the average were removed. We then used Louvain’s algorithm [8] to detect communities within each network based on user interactions. Network analysis was performed using the *NetworkX* [28] and visualizations were made with *Gephi* [29].

It is important to note that the data about individual subreddits are publicly accessible and do not include personally identifiable information (PII). All the data is widely accessible through Reddit’s API. Since our study falls under the category of observational studies conducted on publicly available data, the University Institutional Review Board (IRB) deemed the study to be exempt. This study’s code and edgelist data can be found at <https://github.com/social-nlp-lab/reddit-network>.

D. Topic Modeling

To investigate the characteristics of each community within the ego networks, we performed topic modeling on the collections of subreddit descriptions per identified community. Subreddit descriptions are written by the moderators, thus providing valuable insights into the community’s focus and interests. This process helped us gain a deeper understanding of the topics that are prevalent in each community and allowed us to uncover the unique characteristics of each group.

For topic modeling, we employed the Non-negative Matrix Factorization (NMF) algorithm [30], which proved more efficient for analyzing short texts, like subreddit descriptions, than other models like Latent Dirichlet Allocation (LDA) [31]. We manually examined the topics for each of the communities within each ego network and evaluated their relevance to the corresponding community to determine the dominant topic.

E. Positionality Statement

As shown in previous studies, researchers’ relevant experiences and beliefs can impact research design and data interpretation [32]. All authors of this paper have had previous experience studying drugs in different capacities. All authors also have experience living in an urban city where substance use, and abuse, are prevalent. However, the authors are not part of the community of interest in this study, as none of them have personal experience with substance use disorders.

IV. RESULTS

A. Thematic Analysis

Using the updated coding scheme, the authors annotated the 131 subreddit descriptions. Table I is a summary of subreddits within each of the six categories (i.e., Drug Information, Harm Reduction, Continued Drug Use, Stop/Reduce Drug Use, Experience Seeking, and Miscellaneous). With a subscriber count of over 1.8 million, the most active community is *r/trees*, which serves as a platform for discussions centered around cannabinoids. *r/Ecinternational* had the lowest number of subscribers, a community for followers of the Energy Control International Organization, a drug testing service. When considering the number of subscribers as an indicator of overall interest, it appears that the Continued Drug Use subreddits have a significantly larger subscriber base compared to others. This category has the highest percentage of public subreddits (Restricted? = FALSE) by approximately 6%, however, this openness may not fully explain the significant difference in the average number of subscribers. It was also found that almost 94% of the communities focused on Continued Drug Use do not have an age restriction (“Over 18 community?” = FALSE). The subreddits that received the least amount of attention were those focused on Harm Reduction, excluding the Miscellaneous category. Future work will focus on investigating whether this trend has remained consistent over time, or if there has been a recent increase/decrease in their membership or users’ interest. Currently, the available data is insufficient to make a definitive assessment in this regard.

B. Network Analysis

The network descriptions and top nodes of the four selected subreddits within the Continued Drug Use category

TABLE II: Average centrality measures for each ego network. Networks are pruned on nodes having a degree greater than or equal to the average degree. The top five nodes are presented per centrality measure; excluding the primary node.

Subreddit	Network Descriptives						Top Nodes			
	#Nodes	#Edges	Avg Degree	Avg Cluster	Avg Short Path	#Community	Betweenness	Closeness	Degree	Eigenvector
<i>benzodiazepines</i>	1,134	125,445	221.2	0.76	1.80	7	<i>Drugs, opiates, AskReddit, depression, weed</i>	<i>Drugs, opiates, AskReddit, depression, weed</i>	<i>Drugs, opiates, AskReddit, depression, weed</i>	<i>Drugs, opiates, AskReddit, depression, MDMA</i>
<i>cocaine</i>	1,457	158,015	216.9	0.76	1.85	7	<i>Drugs, AskReddit, opiates, benzodiazepines, LSD</i>	<i>Drugs, AskReddit, opiates, benzodiazepines, LSD</i>	<i>Drugs, AskReddit, opiates, benzodiazepines, LSD</i>	<i>Drugs, AskReddit, depression, opiates, Showerthoughts</i>
<i>LSD</i>	1,805	220,710	244.5	0.77	1.86	9	<i>weed, trees, Psychonaut, shrooms, Psychedelics</i>	<i>weed, Psychonaut, shrooms, trees, AskReddit</i>	<i>weed, Psychonaut, shrooms, trees, AskReddit</i>	<i>weed, Psychonaut, shrooms, trees, DMT</i>
<i>opiates</i>	996	96,781	194.3	0.78	1.80	7	<i>LSD, AskReddit, weed, Drugs, benzodiazepines</i>	<i>AskReddit, LSD, weed, Drugs, benzodiazepines</i>	<i>AskReddit, LSD, weed, Drugs, benzodiazepines</i>	<i>AskReddit, weed, LSD, Showerthoughts, AskDocs</i>

(i.e., *r/LSD*, *r/benzodiazepines*, *r/opiates*, and *r/cocaine*) are presented in Table II. These networks represent the interactions between the four drug-related subreddits and other subreddits on the platform based on their users' activity and interactions. For each network, in addition to the basic statistics, we also extracted top nodes with respect to four centrality measures; degree centrality measures the number of connections a node (subreddit) has to other nodes, closeness centrality measures how quickly a node can reach other nodes in the network, betweenness centrality measures how important a node is in connecting different parts of the network, and eigenvector centrality measures how well-connected a node is to other well-connected nodes.

Our analysis showed that the *r/LSD* subreddit network was the largest, with 1,805 nodes and 220,710 edges. The *r/LSD* and *r/benzodiazepines* had the highest average degree centrality, indicating that the subreddits within these networks, and thus users, are highly interconnected. This suggests that these communities share a high level of engagement, with users frequently interacting with each other and exchanging information. All networks had high clustering coefficients, indicating that the subreddits within each network were closely connected to each other. This suggests that the subreddits within each network form tightly-knit communities, and users within each network are likely to interact with each other and share information. This is likely due to a number of factors, including a shared interest in topics of discussion within each of the subreddits and the role Reddit communities play among PWUD. Reddit has been shown to provide a safe space for PWUD, who often face stigma and isolation, which could lead them to seek support from other people who understand and share similar lived experiences [6]. This phenomenon can also be observed through the average shortest paths between the subreddits in each network, resembling small-world networks. While the high clustering coefficient allows for information to be easily shared within the network, the short average path length allows for information to be quickly transmitted to any part of the network [33]. These two features combine to create a network that is both highly interconnected and efficient.

Additionally, several nodes are shared between the networks, including *r/Drugs*, *r/AskReddit*, and *r/weed*. These nodes commonly have the highest betweenness, closeness, eigenvector, and degree centrality, suggesting that they are in particular significant in connecting the users and information spread between these communities. This could be expected as *r/Drugs* is the largest drug-related subreddit with over

4 million subscribers. *r/AskReddit* is the largest subreddit across the entire platform, with over 40 million subscribers. Furthermore, subreddits like *r/depression*, *r/MDMA*, *r/DMT*, and *r/AskDocs* appear as influential nodes in the networks based on eigenvector centrality. Eigenvector centrality is a measure of the influence of a node in a network, calculated by taking into account the centrality of the node's neighbors [34]. This means that these nodes are connected to nodes that are themselves influential. Although we cannot say specifically how subscribers are discussing DMT and MDMA within these communities, these substances are psychedelic drugs that people often take with a sedative like benzodiazepines to get relief from toxicity symptoms, like muscle spasms [35]. *AskDocs* is a subreddit for people who have questions about medical topics. It can be a valuable resource for people who have questions about drug use, addiction, and recovery, though we cannot know for certain without performing a deeper analysis on the content of the posts.

C. Community Detection

Using Louvain's algorithm [8], we detected communities within each subreddit's ego network based on user interactions. The output of community detection algorithms can be interpreted as a set of subreddits that are more closely related to each other than to other subreddits in the network [8]. Table II shows the number of communities detected in each network. Each community that was identified represents a cluster of subreddits that share common themes, topics, or interests. These subreddits are more likely to be visited by the same group of users, who share similar preferences and are interested in engaging with similar types of content. Initial observation found distinct foci for different communities.

Figure 2 (2a, 2b, 2c, and 2d) show the ego networks of each primary subreddit. In these figures, nodes are subreddits and edges are weighted by the number of users that are shared between the nodes (subreddits). The colors of the edges represent distinct communities within each network, and the size of the node and text reflects betweenness centrality.

D. Topic Modeling

After identifying communities within each network, we analyzed the descriptions of each subreddit inside these communities to extract their themes and relationships using NMF topic modeling [30]. We identified a total of 5 topic clusters for each community in every network. We used these topics to analyze the nature of discussions that are permitted within these subreddits and the characteristics of each community.

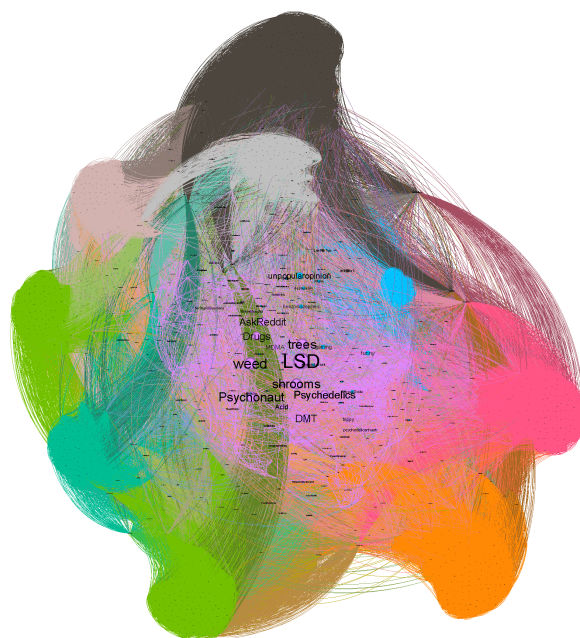
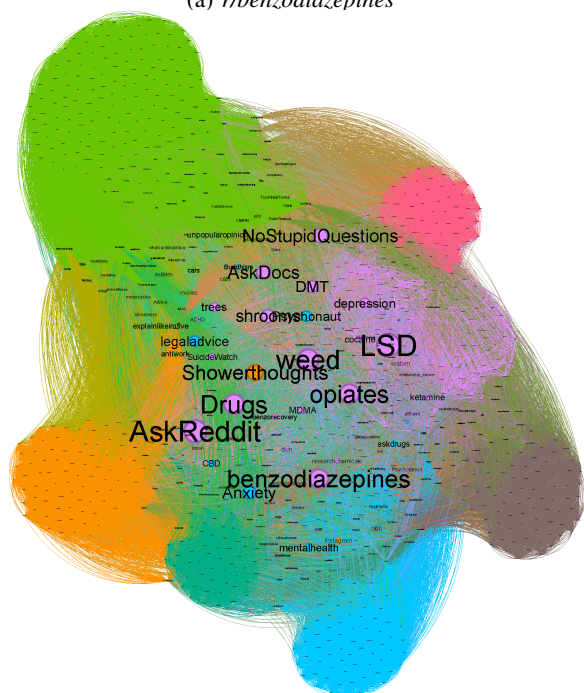
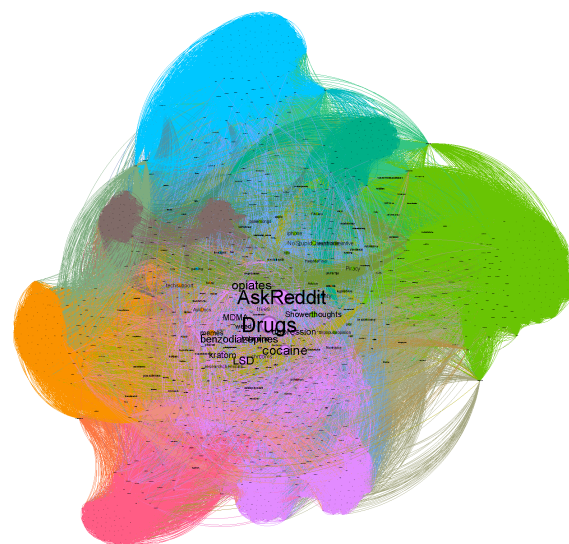
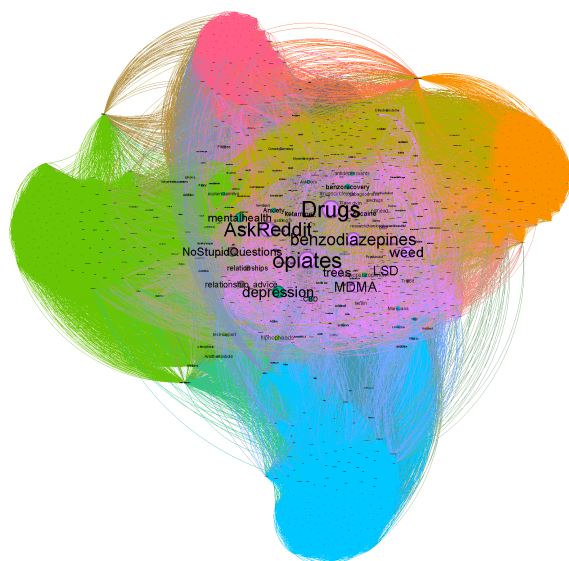


Fig. 2: Ego network of active users in four selected drug subreddits. Nodes: subreddits, Edge weights: number of users between subreddits, Colors: communities, Text size: betweenness centrality.

Across all communities and networks, the most common topic was associated with harm reduction followed by sharing and discussing drug experiences, arts and music, and medical advice. While, the thematic analysis of 131 drug-related subreddits showed that there are not many subreddits dedicated to harm reduction (Table I), finding this topic as one of the most salient topics across all networks suggests a need for more online spaces where people can discuss strategies for reducing harm and get support from others with the same experience. For each network, the most salient topics are as follows: *r/cocaine*: Medical advice, harm reduction, and

sharing experiences and stories (e.g. *r/Drugs*); *r/LSD*: Harm reduction, Q&A based discussions on drug-related support, experiences, and harm (e.g. *r/Drugs*); *r/opiates*: Dedicated discussions on arts, music, and gaming (e.g. *r/gaming*), and Q&A type discussions on personal advice, safe experiences, and stories (e.g. *r/AskReddit*); *r/benzodiazepines*: Dedicated discussions to Q&A legal, medical, personal advice (e.g. *r/NoStupidQuestions*) and peer support for people struggling with mental health disorders (e.g. *r/depression*). These findings reinforce the diverse yet overlapping nature of Reddit users who subscribe to these networks of drug-based discussions.

V. DISCUSSION

Drug-focused subreddits have emerged as thriving online platforms, offering individuals the space to share their personal experiences, seek answers to their inquiries, and find support related to drug use. To gain a comprehensive understanding of these digital communities, we adopted a holistic approach that focused on the larger network as opposed to individual subreddits. Our methodology employed network analysis, community detection, and topic modeling to delve into the nuances of these drug-centric subreddits, both internally (RQ1) and in the broader Reddit landscape (RQ2).

A thematic analysis of 131 drug-related subreddits revealed that the majority of these communities ($N = 61$) are primarily focused on discussions related to Continued Drug Use. Expanding our analysis beyond the drug-specific subreddits and examining users' interaction networks, we observed that despite the limited number of subreddits dedicated to the theme of Harm Reduction, this topic emerged as one of the most prevalent themes within the subreddit descriptions of the broader network. This finding suggests a strong interest in and a need for harm reduction information and dedicated support within the drug-related online communities on Reddit. Future studies should consider the potential for fostering and promoting these specialized communities.

The network analysis revealed three influential subreddits within the networks: *r/AskReddit*, *r/Drugs*, and *r/weed*. Prior research has demonstrated the utility of *r/Drugs*, a community (of almost 4 million subscribers) dedicated to the discussion of all drugs with an emphasis on harm reduction, to be useful as a predictor for future trends in new psychoactive substances (NPS) related exposures [36]. Comparing the timelines of NPS mentions on Reddit with comprehensive toxicology data showed that peak mentions generally occurred before peak exposure, indicating its predictive value. This is important given the small-world nature of our networks, which can be advantageous for accessing information and community support but also risky due to potential misinformation spread, especially in controversial drug-use discussions. Small-world networks have the potential to create homogeneous and polarizing communities [37] that can ultimately affect social attitudes towards PWUD and create more barriers to those suffering from substance use disorder.

Furthermore, subreddits such as *r/Drugs* predominantly serve as a platform for sharing action-facilitating support, assisting individuals in managing the uncertainty associated with substance use [38]. This kind of support can be especially valuable for those seeking emotional and informational assistance during their recovery journey. Within our networks, recovery-oriented subreddits emerged as central nodes with high betweenness centrality. Notably, *r/OpiatesRecovery* and *r/benzorecovery* were prominent counterparts within the *r/opiates* and *r/benzodiazepines* networks, respectively. These subreddits provide a place for users to ask for advice, offer support, and discuss recovery. This trend resonates with broader literature that underscores the significance of online

communities in health and recovery narratives [6]. These communities serve as vital resources for individuals facing challenges. Additionally, the centrality of these recovery-focused subreddits reinforces the established understanding that digital spaces play an instrumental role in fostering peer support and disseminating information in times of vulnerability [39].

VI. CONCLUSION, LIMITATIONS, AND FUTURE WORK

Through a combination of network analysis, community detection, and topic modeling, we studied drug-related discussions on Reddit, one of the most popular social media sites in the US. Our exploration of drug-focused subreddits has unveiled a dynamic online ecosystem. Our results showed that these platforms not only reflect the diverse needs and interests of their users but also hold the potential to shape the discourse surrounding substance use, harm reduction, and recovery. These communities provide a space for people to share their experiences, ask questions, and receive support.

Our study has several limitations. First, the initial list of drug-related subreddits used in this study was created around five years ago and may not be comprehensive. As the digital landscape is continually evolving, newer platforms or subreddits may have emerged that we may not have included in our research. In addition, we restricted our analysis to English-language subreddits, which inherently exclude a vast array of discussions and narratives taking place in other languages. There are subreddits focused on community needs in languages other than English (e.g., *GermanTrees*), and our current analysis does not capture the perspectives, or challenges discussed in those forums. Furthermore, we did not take into consideration the legal classifications of drugs. For example, subreddits about legal drugs, such as alcohol and caffeine, may have different norms and cultures than subreddits about illegal drugs, such as LSD and heroin. Finally, our reliance on descriptions provided by subreddit moderators may not accurately capture the true essence and depth of discussions within these communities. Such descriptions, while useful, are inherently subjective and might not cover the full range of topics or sentiments expressed by the community members. As a result, there might be aspects of these subreddits that our study missed or did not fully represent.

Future research will extend beyond the analysis of Reddit descriptions, delving into users' conversations within the subreddits to gain a more nuanced understanding of the characteristics unique to each subreddit concerning drug-related discussions. Furthermore, we plan to employ our extended coding scheme to broader datasets and communities, evaluating whether conversations and interactions within subreddits align with the intended objectives outlined in moderator-written descriptions. Additionally, the networks could be studied temporally to assess the effects of time and socio-cultural factors, like policy, on the flow of user activity.

REFERENCES

- [1] O. T. Hall, C. Trimble, S. Garcia, P. Entrup, M. Deaner, and J. Teater, "Unintentional drug overdose mortality in years of life lost among

- adolescents and young people in the us from 2015 to 2019,” *JAMA pediatrics*, vol. 176, no. 4, pp. 415–417, 2022.
- [2] S. C. S. R. Center, “Preventing addiction: The importance of drug prevention.” Available at <https://socalsunrise.com/the-importance-of-drug-prevention/>.
 - [3] E. Naserianhanzaei and M. Koschate-Reis, “Effects of Substance Use, Recovery, and Non-Drug-Related Online Community Participation on the Risk of a Use Episode During Remission From Opioid Use Disorder: Longitudinal Observational Study,” *Journal of Medical Internet Research*, vol. 24, p. e36555, Aug. 2022. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Canada.
 - [4] W. Kepner, M. C. Meacham, and A. L. Nobles, “Types and Sources of Stigma on Opioid Use Treatment and Recovery Communities on Reddit,” *Substance Use & Misuse*, vol. 57, pp. 1511–1522, Aug. 2022. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/10826084.2022.2091786>.
 - [5] A. Spadaro, A. Sarker, W. Hogg-Bremer, J. S. Love, N. O’Donnell, L. S. Nelson, and J. Perrone, “Reddit discussions about buprenorphine associated precipitated withdrawal in the era of fentanyl,” *Clinical Toxicology (Philadelphia, Pa.)*, vol. 60, pp. 694–701, June 2022.
 - [6] A. R. D’Agostino, A. R. Optican, S. J. Sowles, M. J. Krauss, K. Escobar Lee, and P. A. Cavazos-Rehg, “Social networking online to recover from opioid use disorder: A study of community interactions,” *Drug and alcohol dependence*, vol. 181, pp. 5–10, Dec. 2017.
 - [7] NCADS, “Substance abuse and addiction statistics,” 2023. <https://drugabusestatistics.org/>.
 - [8] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, p. P10008, Oct. 2008. arXiv:0803.0476 [cond-mat, physics:physics].
 - [9] S. Atske, “Social Media Use in 2021,” Apr. 2021.
 - [10] M. J. Barratt, “Discussing illicit drugs in public internet forums: visibility, stigma, and pseudonymity,” in *Proceedings of the 5th International Conference on Communities and Technologies, C&T ’11*, (New York, NY, USA), pp. 159–168, Association for Computing Machinery, June 2011.
 - [11] A. Preiss, P. Baumgartner, M. J. Edlund, and G. V. Bobashev, “Using Named Entity Recognition to Identify Substances Used in the Self-medication of Opioid Withdrawal: Natural Language Processing Study of Reddit Data,” *JMIR Formative Research*, vol. 6, p. e33919, Mar. 2022. Company: JMIR Formative Research Distributor: JMIR Formative Research Institution: JMIR Formative Research Label: JMIR Formative Research Publisher: JMIR Publications Inc., Toronto, Canada.
 - [12] L. Bouzoubaa, R. Ehsani, P. Chatterjee, and R. Rezapour, “The evolution of substance use coverage in the philadelphia inquirer,” *arXiv preprint arXiv:2307.01299*, 2023.
 - [13] S. Nielsen and M. J. Barratt, “Prescription drug misuse: is technology friend or foe?,” *Drug and Alcohol Review*, vol. 28, pp. 81–86, Jan. 2009.
 - [14] Z. Davey, F. Schifano, O. Corazza, P. Deluca, and Psychonaut Web Mapping Group, “e-Psychonauts: conducting research in online drug forum communities,” *Journal of Mental Health (Abingdon, England)*, vol. 21, pp. 386–394, Aug. 2012.
 - [15] R. L. Graves, J. Perrone, M. A. Al-Garadi, Y.-C. Yang, J. Love, K. O’Connor, G. Gonzalez-Hernandez, and A. Sarker, “Thematic Analysis of Reddit Content About Buprenorphine-naloxone Using Manual Annotation and Natural Language Processing Techniques,” *Journal of Addiction Medicine*, vol. 16, no. 4, pp. 454–460, 2022.
 - [16] F. Schifano, P. Deluca, A. Baldacchino, T. Peltoniemi, N. Scherbaum, M. Torrens, M. Farré, I. Flores, M. Rossi, D. Eastwood, C. Guionnet, S. Rawaf, L. Agosti, L. Di Furia, R. Brigada, A. Majava, H. Siemann, M. Leoni, A. Tomasin, F. Rovetto, and A. H. Ghodse, “Drugs on the web; the Psychonaut 2002 EU project,” *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 30, pp. 640–646, June 2006.
 - [17] J. Pestana, F. Beccaria, and E. Pettrilli, “Psychedelic substance use in the Reddit psychonaut community. A qualitative study on motives and modalities,” *Drugs and Alcohol Today*, vol. 21, pp. 112–123, Jan. 2020. Publisher: Emerald Publishing Limited.
 - [18] A. M. Bunting, D. Frank, J. Arshonsky, M. A. Bragg, S. R. Friedman, and N. Krawczyk, “Socially-supportive norms and mutual aid of people who use opioids: An analysis of reddit during the initial covid-19 pandemic,” *Drug and alcohol dependence*, vol. 222, p. 108672, 2021.
 - [19] D. Esty and R. Rushing, “The promise of data-driven policymaking,” *Issues in Science and Technology*, vol. 23, no. 4, 2007.
 - [20] S. Pandrekar, X. Chen, G. Gopalkrishna, A. Srivastava, M. Saltz, J. Saltz, and F. Wang, “Social Media Based Analysis of Opioid Epidemic Using Reddit,” *AMIA Annual Symposium Proceedings*, vol. 2018, pp. 867–876, Dec. 2018.
 - [21] K. B. Rhidenour, K. Blackburn, A. K. Barrett, and S. Taylor, “Mediating medical marijuana: exploring how veterans discuss their stigmatized substance use on reddit,” *Health Communication*, vol. 37, no. 10, pp. 1305–1315, 2022.
 - [22] O. Enghoff and J. Aldridge, “The value of unsolicited online data in drug policy research,” *International Journal of Drug Policy*, vol. 73, pp. 210–218, Nov. 2019.
 - [23] L. Dinh, R. Rezapour, L. Jiang, and J. Diesner, “Enhancing structural balance theory and measurement to analyze signed digraphs of real-world social networks,” *Frontiers in Human Dynamics*, vol. 4, p. 1028393, 2023.
 - [24] S. H. Cook, J. A. Bauermeister, D. Gordon-Messer, and M. A. Zimmerman, “Online Network Influences on Emerging Adults’ Alcohol and Drug Use,” *Journal of Youth and Adolescence*, vol. 42, pp. 1674–1686, Nov. 2013.
 - [25] M. Copeland, J. C. Fisher, J. Moody, and M. E. Feinberg, “Different Kinds of Lonely: Dimensions of Isolation and Substance Use in Adolescence,” *Journal of Youth and Adolescence*, vol. 47, pp. 1755–1770, Aug. 2018.
 - [26] Reddit User, “List of subreddits about drugs, addiction, drug abuse, etc.,” 2023. [Online; accessed 17-October-2023, www.reddit.com/r/ListOfSubreddits/comments/abmkx6/list_of_subreddits_about_drugs_addiction_drug/].
 - [27] L. Marston, *Introductory Statistics for Health and Nursing Using SPSS*. London: SAGE Publications Ltd, 2010.
 - [28] A. A. Hagberg, D. A. Schult, and P. J. Swart, “Exploring Network Structure, Dynamics, and Function using NetworkX,” in *Proceedings of the 7th Python in Science Conference* (G. Varoquaux, T. Vaught, and J. Millman, eds.), (Pasadena, CA USA), pp. 11 – 15, 2008.
 - [29] M. Bastian, S. Heymann, and M. Jacomy, “Gephi: An open source software for exploring and manipulating networks,” 2009.
 - [30] V. P. Pauca, F. Shahnaz, M. W. Berry, and R. J. Plemmons, “Text mining using non-negative matrix factorizations,” in *Proceedings of the 2004 SIAM international conference on data mining*, pp. 452–456, SIAM, 2004.
 - [31] D. M. Blei, “Probabilistic topic models,” *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, 2012.
 - [32] S. Secules, C. McCall, J. A. Mejia, C. Beebe, A. S. Masters, M. L. Sánchez-Peña, and M. Svyantek, “Positionality practices and dimensions of impact on equity research: A collaborative inquiry and call to the community,” *Journal of Engineering Education*, vol. 110, no. 1, pp. 19–43, 2021.
 - [33] M. D. Humphries and K. Gurney, “Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence,” *PLOS ONE*, vol. 3, p. e0002051, Apr. 2008. Publisher: Public Library of Science.
 - [34] J. Golbeck, “Chapter 3 - Network Structure and Measures,” in *Analyzing the Social Web* (J. Golbeck, ed.), pp. 25–44, Boston: Morgan Kaufmann, Jan. 2013.
 - [35] D. E. Nichols, “Psychedelics,” *Pharmacological Reviews*, vol. 68, pp. 264–355, Apr. 2016.
 - [36] J. Lu, S. Sridhar, R. Pandey, M. A. Hasan, and G. Mohler, “Investigate transitions into drug addiction through text mining of reddit data,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2367–2375, 2019.
 - [37] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociochi, “The spreading of misinformation online,” *Proceedings of the National Academy of Sciences*, vol. 113, pp. 554–559, Jan. 2016. Publisher: Proceedings of the National Academy of Sciences.
 - [38] K. Wombacher, S. E. Sheff, and N. Itrich, “Social support for active substance users: A content analysis of r/drugs,” *Health Communication*, vol. 35, p. 756–765, May 2020.
 - [39] K. Latha, K. S. Meena, M. R. Pravitha, M. Dasgupta, and S. K. Chaturvedi, “Effective use of social media platforms for promotion of mental health awareness,” *Journal of Education and Health Promotion*, vol. 9, p. 124, May 2020.