

# Exploring Musical, Lyrical, and Network Dimensions of Music Sharing Among Depression Individuals

Qihan Wang  
qihan.wang@aalto.fi  
Aalto University  
Espoo, Finland

Zeyad Alghamdi  
zalgham1@asu.edu  
Arizona State University  
Tempe, Arizona, USA

Anique Tahir  
artahir@asu.edu  
Arizona State University  
Tempe, Arizona, USA

Huan Liu  
huanliu@asu.edu  
Arizona State University  
Tempe, Arizona, USA

## ABSTRACT

Depression has emerged as a significant mental health concern due to a variety of factors, reflecting broader societal and individual challenges. Within the digital era, social media has become an important platform for individuals navigating through depression, enabling them to express their emotional and mental states through various mediums, notably music. Specifically, their music preferences, manifested through sharing practices, inadvertently offer a glimpse into their psychological and emotional landscapes. This work seeks to study the differences in music preferences between individuals diagnosed with depression and non-diagnosed individuals, exploring numerous facets of music, including musical features, lyrics, and musical networks. The music preferences of individuals with depression through music sharing on social media, reveal notable differences in musical features and topics and language use of lyrics compared to non-depressed individuals. We find the network information enhances understanding of the link between music listening patterns. The result highlights a potential echo-chamber effect, where depression individual's musical choices may inadvertently perpetuate depressive moods and emotions. In sum, this study underscores the significance of examining music's various aspects to grasp its relationship with mental health, offering insights for personalized music interventions and recommendation algorithms that could benefit individuals with depression.

## ACM Reference Format:

Qihan Wang, Anique Tahir, Zeyad Alghamdi, and Huan Liu. 2018. Exploring Musical, Lyrical, and Network Dimensions of Music Sharing Among Depression Individuals. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX) ACM, New York, NY, USA, 9 pages. <https://doi.org/XXXXXXX.XXXXXXX>*

---

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 ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Conference acronym 'XX, June 03–05, 2018, Woodstock, NY*

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXXX>

## 1 INTRODUCTION

Depression is a prevalent mental disorder characterized by persistent feelings of sadness, a diminished interest, or a lack of pleasure in daily activities [24]. Depression impacts approximately 3.8% of the global population across all ages, genders, and cultural groups [37], posing a significant worldwide challenge [53]. Distinct from regular mood fluctuations, depression is characterized by its prolonged duration, and may manifest in feelings of excessive guilt or low self-worth, loss of interest or pleasure, hopelessness about the future, disturbed sleep or appetite, and poor concentration [19, 55]. With the development of social media, individuals experiencing depression tend to express their thoughts and share their feelings with their peers and audiences online [4]. The shared content includes a variety of forms, such as text, photos, news and music.

Music, intertwining with daily lives, enables individuals to regulate emotions, maintain interpersonal relationships, and express their identities [31, 35]. As people connect their own memories and thoughts with music [34], music sharing can serve as a mirror to reflect people's moods and feelings towards life events [25, 48]. Building on this, individuals with depression may perceive music differently and exhibit distinct musical preferences. Previous psychology research has shown that individuals with a tendency towards depression demonstrated a preference for sad music [23], and individuals who are diagnosed by depression perceive music as conveying more negatively emotion [17]. However, most of the previous research on music and depression focused on the emotion of the songs, ignoring other musical features that can be quantified or measured by rhythmic and structural characteristics, such as tempo, energy, mode and acousticness. In addition, they often overlook elements beyond musical features, such as lyrics and associated networks.

In sum, while prior investigations have primarily studied the emotional aspects of songs, the broader spectrum of musical attributes, lyrics, and associated networks largely remains unexplored. This significant gap presents a chance to explore in more detail the differences of music preferences between individuals with and without depression. Given this context, our study takes a broad approach, carefully examining not just various musical features, but also delving into lyrics and network characteristics among social media users. Based on a publicly available dataset [2] that encompasses music sharing data from users who have self-expressed that they have been diagnosed with depression, together with data from

randomly-selected control users (non-depression users), we studied 4 research questions:

**RQ1: Do musical features of music sharing differ between depression and non-depression individuals?** By collecting musical features of each song from Spotify API [3], we discern any notable variations between the depression and non-depression cohorts. We found that songs shared by depression are significant higher in energy, mode, acousticness, instrumentalness, and duration, and lower in speechiness, time signature, popularity and danceability.

**RQ2: What are specific artists or genres that are commonly shared by depression and non-depression individuals?** We presented qualitative visualizations for artists and genres shared by depression groups and control groups, and leveraged TF-IDF to present artists and genres that are uniquely shared by depression groups.

**RQ3: Do topics and language use in lyrics of music sharing differ between depression and non-depression individuals?** We conducted a multi-faceted lyric level analysis. Our result showed that music shared by individuals with depression have higher proportion in the topic of emotional reflection and religious language, yet lower proportion in the topic of explicit (RAP) languages and personal relationships. Furthermore, individuals with depression prefer lyrics with shorter sentences.

**RQ4: What are the network relationships regarding music shared by depression and non-depression individuals?** Based on Spotify Playlist Dataset [1], we categorized playlists as depression/control playlists by the amount of songs that shared by depression users and conducted community detection for all playlists. Our result indicated that the depression playlists cluster together, suggesting a network connection between depression songs. Combined and coherent with the previous lyric analysis, we found that emotional reflection and religious or biblical language are dominant topics in depression playlist communities.

In synthesizing our findings, we discern crucial implications across two pivotal dimensions. Firstly, our results demonstrated social and cognitive characteristic features of individuals with depression in relation to their music preferences. Secondly, we delved into the consequences engendered by these specific musical preferences. We found that the music preference of individuals with depression potentially encapsulates users within a depression Echo-Chamber Effect, where the music exacerbates prevailing depressive moods and emotions. This explorative analysis not only heightened our understanding of the nuanced relationships between music preferences and mental health but also illuminated pathways for the development of music recommendation algorithms that were attuned to the needs of individuals navigating mental health disorders.

## 2 RELATED WORK

Music, an integral component of human interaction, has seen its influence magnified due to advances in digital technology, especially on platforms like social media [32, 33]. These platforms have become central stages where music is shared, offering insights into individuals' identity and self-presentation [31, 50]. In this work,

we focus on its relationship with mental states, particularly depression. Specifically, we explore the multifaceted connections between music sharing and depression, including musical features, lyrics, and social network relationships. This section aims to provide an overview, drawing from existing research to shape our discourse on these interconnected themes.

### 2.1 Music Features and Its Multifaceted Connection to Individuals

Our study aims to incorporate a diverse array of musical features, providing a more comprehensive understanding of the music preferences of individuals with depression. Spanning from rhythmic and structural to timbral characteristics, these features serve as quantitative representations of various dimensions of music, enabling a wide range of applications such as music recommendation, playlist generation, and exploratory musicological research [3]. These features encompass various attributes such as danceability, duration, energy, instrumentalness, liveness, loudness, speechiness, tempo, and valence [6, 26]. Each feature encapsulates a distinct aspect of a song, elucidating its various musical and acoustical properties; for instance, danceability refers to how suitable a track is for dancing, often through examining elements like beat regularity and strength, and valence serves to describe the musical positiveness conveyed by a track, distinguishing between sonically positive (e.g., happy, cheerful) and negative (e.g., sad, depressive) emotional expressions. Moreover, research have found that there's a notable correlation between these musical features and individuals' traits: individual's personality aligns with their affinity or preference for specific musical features, including mode, tempo, and dynamics [20, 21].

### 2.2 Overview of Depression and Music

Music preference is intricately linked with an individual's conditions and traits, including mental health conditions [36], personality [16], and intelligence [10]. Specifically, when examining the connection with depression, past research has predominantly centered on musical genres and sentiments. In the context of genres, prior studies suggest that listening to soul music (hip hop and R&B) is a predictor of lower depression levels [41]. As for sentiment, previous research has demonstrated that individuals diagnosed with depression perceive music as more negatively valence [17]. Such individuals exhibit a tendency to gravitate towards negatively valenced music, mirroring their chronic mood state [54]. Moreover, listening to sad music can significantly increase the depressed feelings of individuals with depression [23].

Nevertheless, music has also been identified as a therapeutic tool for individuals with depression. Studies indicate that music listening can lead to a notable reduction in depression levels [40]. This therapeutic approach, termed music therapy, has been corroborated as an efficacious treatment for depression [18].

### 2.3 Overview of Depression and Social Media

Recent research increasingly harnesses user-generated content on social media to explore mental health-related issues, generally employing two primary approaches: (1) analyzing the behavior patterns, cognition, and emotional aspects of mental health disorders

[56], and (2) developing automatic detection systems for mental health problems or anomalies [15, 28]. These studies broadly utilizes existing tools and methods from data mining, natural language processing, social network analysis, and machine learning. Within this framework, a spectrum of mental health disorders, including major depressive disorders [15], eating disorders [22], and ADHD [27], have been explored. Our work is anchored in the first approach, and it aims to discern the variance in music-sharing behaviors between individuals diagnosed with depression and those who aren't.

Building upon the foundation of the first approach, earlier research focusing on depression and its manifestation on social media has delved into aspects like social interaction, linguistic expression, and emotional displays of affected users. From a sociological standpoint, it's been observed that individuals with depression tend to manifest a decrease in social activity and form highly clustered ego-networks [14]. From linguistical perspectives, it has been shown that individuals with depression express more distorted thinking, and their online language is generally deemed depressogenic [7]. Notably, the linguistic patterns and topic-centric vocabulary of depressed individuals significantly deviate from their non-depressed counterparts [47]. On the emotional front, it's been found that self-reported depression is related to the usage of personal and negative pronoun language categories [38]. Furthermore, individuals with depression form more negative sentences from scrambled words. In other words, they tend to form more pessimistic sentences than non-depressed participants, even when they are not in a negative mood [52]. Moreover, from topic perspectives, people with depression have amplified relational and medicinal concerns and express greater religious involvement [14]. While previous research has established a discrepancy in the behavior and expression between individuals with and without depression on social media, spanning sociality, language expression, and emotion, a gap persists in understanding the domain of music preferences. Our study aspires to bridge this gap, aiming to discern how music sharing practices diverge between these two groups, consequently presenting a fresh lens to the existing discourse.

### 3 EVALUATION & RESULTS

In this section, we present the findings obtained from our analysis focusing on three primary domains: musical variables, lyrics and network. Our objectives were to: (1) identify significant differences in musical variables, artists and genres between depression and non-depression groups by utilizing statistical analysis (2) present artist and genres differences qualitatively by word cloud (3) understand the difference of topics and language use of the lyrics between depression and non-depression groups by LDA and LIWC (4) explore the connection of depression playlists by leveraging community detection.

#### 3.1 Data Collection

To achieve our research objectives, we carefully selected a dataset that aligns with the intersection of mental health and musical engagement, which is a crucial part of our investigation. But also aiding us with a notable advantage, that is, the inclusion of a depression and a control group, which provides a foundational comparison and paves the way for a deeper and more comprehensive analysis

of the interplay between music-sharing behaviors and mental states in a digital context. Therefore, we sourced an anonymized publicly available dataset that encompasses social media information [2]. In our selected dataset, the 'Depression Group' is characterized by users who have self-identified with a depression diagnosis on Twitter, while the 'Control Group', also known as 'Non-Depression Group', includes those who never tweeted depression or any other mental disorders. Every entry in this dataset provides specifics like the song's title, artist, and cleaned-up lyrics. There's also a marker to identify whether the song-sharer publicly disclosed a depression diagnosis on Twitter. We further ensured a balanced representation of music-sharing posts between both groups. For a richer analysis of the link between music-sharing and depression, we tapped into the Spotify API [3, 46] to extract musical attributes for each shared piece. This extraction process yielded 14 numerical variables, which are described in table 1. Each variable reveals a unique facet of the musical piece. Given the wide spectrum of genres a song can belong to, the collective attributes present a panoramic perspective on musical styles and their potential implications for understanding an individual's psychological state.

For network analysis, we leverage the Spotify Million Playlist Dataset [1], which contains titles for songs and playlists created by users on the Spotify platform between January 2010 and October 2017. We aimed to examine if playlists with songs linked to depression cluster distinctively compared to those without such songs, aiming to shed light on the intricate relationship between music preferences and mental health.

#### 3.2 Musical Features

In addressing RQ1 through our exploration of musical features, we aimed to identify any significant differences in musical variables between depression and non-depression groups. Our null hypothesis, denoted as  $H_0$ , posits no significant difference in the evaluated musical attributes between the two cohorts. We subjected this hypothesis to empirical verification employing a suite of statistical tests. We found that there are significant differences between depression and control group in danceability ( $MD = -0.041, p < 0.001$ ), energy ( $MD = -0.012, p < 0.001$ ), speechiness ( $MD = -0.030, p < 0.001$ ), acousticness ( $MD = 0.009, p < 0.05$ ), instrumentalness ( $MD = 0.022, p < 0.001$ ), Time Signature ( $MD = -0.026, p < 0.001$ ), popularity ( $MD = -1.000, p < 0.001$ ). Moreover, the duration of the song is significantly longer in depression group ( $MD = 8875.318, p < 0.001$ ). On the other hand, there are significantly more major mode and less minor mode in depression group ( $MD = 0.044, p < 0.001$ ). Additionally, there is no difference in key, loudness, liveness, valence, and tempo.

In addition to the quantitative analysis for artists and genres for answering RQ2, we also employ qualitative visualizations to intuitively represent the prevalent elements in our dataset.

Specifically, we construct word clouds of depression and control group for artists (figure 1a and 1b) and genres (figure 1d and 1e). By leveraging TF-IDF, we present the unique artist 1c and genres 1f in depression group.

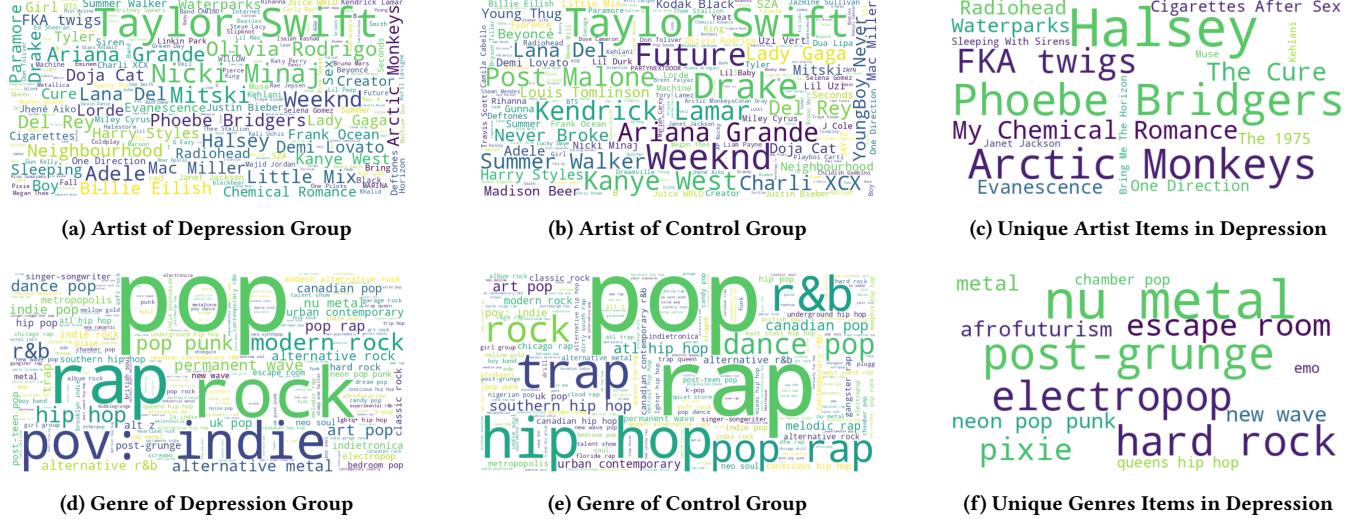


Figure 1: Qualitative Visualization of Artist and Genres

### 3.3 Lyrics Analysis

Although the analysis of musical features have provided invaluable insights, there remains an intricate layer of understanding that can only be unraveled through an examination of lyrics. Lyrics, imbued with semantic and emotional content, serve as a conduit for artists to express intricate sentiments, and for listeners to resonate with these expressions. The lyrical content can amplify, complement, or even contrast the emotional tenor conveyed by musical elements. Given that individuals with depression may be particularly attuned to the emotional and thematic content of lyrics (add citation), a granular analysis at the lyric level is paramount. It facilitates a more comprehensive understanding of the specific themes, narratives, and emotional expressions that resonate with this demographic. This, in turn, could unveil nuanced patterns and correlations, offering a richer, multidimensional perspective on the intricate interplay between music, lyrics, and mental health, and potentially informing targeted, lyrics-based interventions for individuals with depression.

In pursuit of **RQ3** that focusing on the lyrics differences, we conducted multi-faceted lyric level analysis. We use traditional topic modelling to distinguish clusters at a linguistic level. In addition, we use LIWC based approach to analyze the lyrics at a more emotional, cognitive, and structural level.

**3.3.1 Topic Modeling.** In our lyrics analysis, we employ Latent Dirichlet Allocation (LDA) to identify latent topics from the corpus of song lyrics, aiming to unveil thematic distinctions between the depression and control groups. LDA is a generative probabilistic model in which each document (in this case, a set of lyrics) is represented as a random mixture of latent topics [9]. Each topic is characterized by a distribution over words in the vocabulary.

Given a corpus of lyrics  $D$ , each song's lyrics  $d \in D$  is modeled as a finite mixture over an underlying set of topics. Each topic  $k$ , represented as  $\beta_k$ , is a distribution over the vocabulary of words. The topics are assumed to have been drawn from a Dirichlet distribution with parameter  $\eta$ :

$$\beta_k \sim \text{Dir}(\eta), \quad k = 1, \dots, K.$$

Each document (or set of lyrics)  $d$  is associated with a vector of topic proportions  $\theta_d$ , which is drawn from a Dirichlet distribution with parameter  $\alpha$ :

$$\theta_d \sim \text{Dir}(\alpha), \quad d = 1, \dots, D.$$

Each word  $w_{di}$  in document  $d$  is generated by first sampling a topic  $z_{di}$  from the multinomial distribution parameterized by  $\theta_d$ , and then sampling the word from the multinomial distribution parameterized by  $\beta_{z_{di}}$ :

$$z_{di} \sim \text{Multinomial}(\theta_d), \quad w_{di} \sim \text{Multinomial}(\beta_{z_{di}}).$$

We use the estimated topic-word distributions and document-topic distributions to analyze the thematic content of lyrics in both the depression and control groups. Statistical tests, such as the Proportion Z-Test, are then applied to evaluate the significance of the observed differences in topic proportions between the two groups. The most common 15 topics in the lyrics are shown in table 3

The Proportion Z-Test is employed to assess whether the proportions of the topic differ significantly between depression and control group. Compared to the control group, there is a statistically significant higher proportion of Music Recording ( $p < 0.05$ ,  $MD = 0.002$ ), Religious or Biblical Language ( $p < 0.01$ ,  $MD = 0.005$ ), Everyday Life ( $p < 0.01$ ,  $MD = 0.0039$ ), Entertainment Industry ( $p < 0.01$ ,  $MD = 0.002$ ), Recollection and Creativity ( $p < 0.05$ ,  $MD = 0.003$ ), Everyday Experiences ( $p < 0.001$ ,  $MD = 0.036$ ), Emotional Reflection ( $p < 0.001$ ,  $MD = 0.093$ ), Everyday Conversations ( $p < 0.01$ ,  $MD = 0.004$ ) in the depression group, and there is a statistically significant lower proportion of Explicit or RAP Language ( $p < 0.001$ ,  $MD = -0.084$ ), Expressive Language ( $p < 0.001$ ,  $MD = -0.0403$ ), Musical Artists and Collaborations ( $p < 0.001$ ,  $MD = -0.007$ ), and

**Table 1: Description of Numerical Variables from Spotify API**

Variable	Description
Danceability	Danceability is a measure from 0.0 to 1.0 which describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity.
Energy	Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy.
Key	The key the track is in, using standard Pitch Class notation.
Loudness	Overall loudness of a song in decibels (dB).
Mode	Modality of a track; 1 represents major and 0 represents minor.
Speechiness	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value.
Acousticness	A confidence measure from 0.0 to 1.0 of whether the track is acoustic.
Instrumentalness	Predicts whether a track contains no vocals.
Liveness	Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live.
Valence	A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track.
Tempo	The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration.
Duration_ms	Duration of the song in milliseconds.
Time_Signature	Estimated overall time signature of a track, which specify how many beats are in each bar
Popularity	Measure of the track's popularity based on number of plays.

Personal Relationships ( $p = 0.01$ ,  $MD = -0.010$ ) in the depression group.

**3.3.2 LIWC Across Single-Occurrence and Repeated Songs.** To explore the characteristics of songs, we employed the Linguistic Inquiry and Word Count (LIWC) tool [12]. This tool provides features that capture the text's linguistic, emotional, cognitive, and structural elements. We designed two subsets of the original dataset, one for each of the depression and control groups, while eliminating overlapping songs to maintain methodological rigor.

For a given lyric  $L$ , LIWC generates a feature vector as follows:

$$F(L) = \begin{bmatrix} f_{\text{linguistic}}(L) \\ f_{\text{emotional}}(L) \\ f_{\text{cognitive}}(L) \\ f_{\text{structural}}(L) \end{bmatrix}$$

**Table 2: Statistical Analysis of Musical Difference**

Variable	T-value	df	MD	p-value
danceability	-17.070585	18262	-0.041	<b>0.000***</b>
energy	4.03491	18262	0.012	<b>0.000***</b>
key	-0.127663	18262	-0.007	0.898
loudness	-2.202757	18262	-0.102	<b>0.028*</b>
mode	6.087954	18262	0.044	<b>0.000***</b>
speechiness	-18.630115	18262	-0.030	<b>0.000***</b>
acousticness	2.229929	18262	0.009	<b>0.026*</b>
instrumentalness	8.837124	18262	0.022	<b>0.000***</b>
liveness	0.815591	18262	0.002	0.415
valence	-0.700019	18262	-0.002	0.484
tempo	1.768637	18262	0.778	0.077
duration_ms	7.613851	18262	8875.318	<b>0.000***</b>
time_signature	-4.905537	18262	-0.026	<b>0.000***</b>
popularity	-3.622495	18262	-1.000	<b>0.000***</b>

\* refers to the significance(\*  $p < 0.05$  , \*\*  $p < 0.01$  , \*\*\*  $p < 0.001$ )

Where:

$f_{\text{linguistic}}$  : Linguistic attributes of  $L$

$f_{\text{emotional}}$  : Emotional tone and sentiment in  $L$

$f_{\text{cognitive}}$  : Cognitive complexity within  $L$

$f_{\text{structural}}$  : Structural and grammatical aspects of  $L$

These feature vectors facilitate nuanced comparisons between the 'control' and 'depression' groups, enriching our understanding of their linguistic and emotional preferences.

The mean of each LIWC feature for both classes was calculated, followed by the computation of the Mean Difference (MD):

$$MD = \frac{1}{m} \sum_{i=1}^m x_{i,\text{Control}} - \frac{1}{n} \sum_{j=1}^n x_{j,\text{Depression}} \quad (1)$$

Here,  $m$  and  $n$  represent the sample sizes of the "control" and "depression" classes, respectively.  $x_{i,\text{Control}}$  and  $x_{j,\text{Depression}}$  are the specific feature values in each class.

Finally, we identified the top 10 features with the most significant mean differences, specifying the directionality to indicate whether the "control" or "depression" group is favored.

- **Word Count (WC)** (+360.38): More complex narratives in the depression group.
- **Words Per Sentence (WPS)** (-14.73): Simpler sentences in the depression group.
- **Clout** (+2.70): Greater social influence in lyrics favored by the depression group.
- **Punctuation (AllPunc)** (-1.46): Less formal writing styles in the depression group.
- **Function Words (Function)** (+1.25): More complex cognitive processing in the depression group.
- **Linguistic Processes (Linguistic)** (+1.05): Intricate language use in the depression group.
- **Dictionary Words (Dic)** (+1.00): Broader vocabulary in the depression group.
- **Comma Usage (Comma)** (-0.84): Less intricate sentences in the depression group.

**Table 3: Topic Modeling Result**

Topic	Interpretation	Most representative topic
0	Music Recording	studio, recorded, nyc, gun, machine
1	Religious or Biblical Language	god, king, thou, said, man
2	Everyday Life	people, going, year, like, school, work
3	Entertainment Industry	taylor, award, music, batman, man
4	Love and Relationships	baby, love, like, want, girl
5	Political or Philosophical Discussion	great, man, people, power, state
6	Conversational Language	say, tell, know, dont, like
7	Explicit or RAP Language	shit, bitch, fuck, nigga, money
8	Recollection and Creativity	remember, love, night, christmas, summer
9	Expressive Language	dont, know, love, like, want
10	Individual Names and Personal Interactions	stephen, zoe, man, old, say
11	Time and Personal Relations	love, january, april, young, july
12	Timeframes and Specific Events	january, march, april, february, june
13	Legal and Constitutional Discourse	state, court, law, power, congress
14	Everyday Experiences	like, come, life, home, hand
15	Musical Artists and Collaborations	feat, lil, west, kanye, travis
16	Emotional Reflection	time, away, heart, love, way
17	Explicit or Medical Words	nigga, health, secretary, bitch, service
18	Personal Relationships	love, like, young, boy, girl
19	Everyday Conversations	like, said, time, day, know

- **Big Words (BigWords)** (+0.75): Preference for complex words in the depression group.
- **Authenticity (Authentic)** (+0.71): Slight preference for authentic lyrics in the depression group.

### 3.4 Network Analysis

To answer **RQ4**, the question regarding network relationships, we analyze music sharing patterns, adopting a network-centric approach to elucidate the underlying structures and associations that may characterize the distinctions between depressed and non-depressed individuals. The core objective is to ascertain whether there exists a notable clustering propensity among songs shared by individuals with depression within distinct communities of music playlists. Our analysis is based on the Million Spotify Playlist Dataset [1].

Playlists were pruned to retain only those songs that have a corresponding presence in our pre-established music-sharing dataset. A playlist is classified as a ‘depression playlist’ if a majority, i.e., over 50%, of its comprising songs have been shared by individuals identified with depression on Twitter. In the constructed network, nodes epitomize individual playlists. Edges, on the other hand, are illustrative of the existence of shared songs between pairs of playlists, with the weight of each edge being indicative of the quantity of shared songs. We used community detection on this network to find potential patterns and associations inherent to the songs shared by individuals with depression.

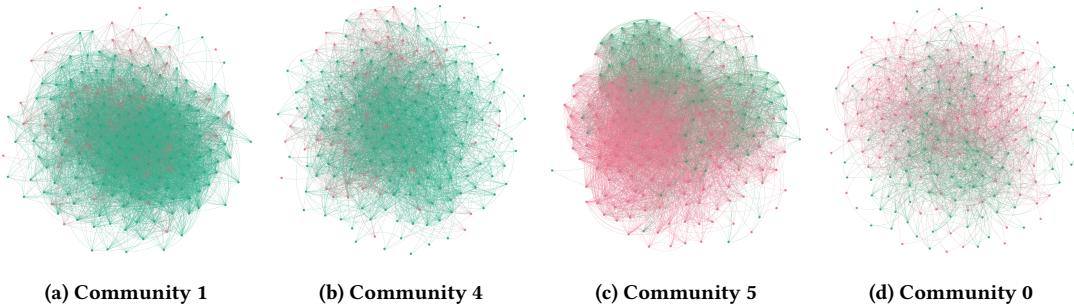
Community detection analysis yielded distinct patterns. It unveiled diverse degrees of dominance by depression and non-depression playlists, indicative of heterogeneous user behaviors and music preferences. For Community detection we used the Louvain algorithm[49]. 40 communities were detected in total. Of these, 9 communities have more than 10 nodes each. Among the 5 communities were

dominated by the depressed population, while 2 communities were dominated by control. 2 communities showed almost equal amounts of depressed and control individuals.

In the ensuing analysis, playlists associated with depression are represented in green, while those unrelated to depression are depicted in pink. Communities 1 and 4 are archetypal of clusters dominant in depression, as illustrated in Figures 2a and 2b. This observation is contrasted by Community 5 (Figure 2c), which is characterized by a dominance of non-depression playlists. Community 0, highlighted in Figure 2d, presents a balanced amalgamation of both depression and non-depression playlists.

The dichotomy in music sharing patterns, as evidenced through the network analysis, provides valuable insights and augments our comprehensive understanding of the behavioral and preferential disparities linked to depression. Each community, delineated by distinct characteristics, affirms the premise that music preferences and sharing behaviors are intrinsically associated with the mental and emotional states of individuals. This investigation not only amplifies the existing body of knowledge but also lays a substantive foundation for subsequent research endeavours aiming at the nexus of mental health, digital behaviors, and music preferences.

We additionally leveraged Latent Dirichlet Allocation (LDA) and community detection to identify the prevailing topic within each community. Intriguingly, communities 3 and 4, which are dominated by depression playlists, were overwhelmingly governed by the Emotional Reflection topic. Similarly, community 8, heavily associated with depression playlists, was primarily dominated by Religious or Biblical Language topic. These insights are consistent with preceding statistical analyses, indicating a higher prevalence of Emotional Reflection and Religious or Biblical Language topics within songs related to depression.



## 4 DISCUSSION

### 4.1 Summary of Findings

In this work, we analyzed the difference of music sharing between depression and non-depression group from the perspectives of musical features, lyrics and musical networks. Statistical analysis of musical features showed that there are songs shared by depression are higher in energy, mode, acousticness, instrumentality, and duration, and lower in speechiness, time signature, popularity and danceability. Lyric analysis showed that depression music sharing have higher proportion in music recording, religious or biblical language, everyday life, entertainment industry, recollection and creativity, everyday experiences, emotional reflection, and everyday conversations topics, yet lower proportion in explicit or RAP language, expressive language, musical artists and collaborations, and personal relationships topics. Additionally, individuals with depression may prefer songs with shorter and simpler sentences, more assertive or confident language. In addition, Network analysis revealed that depression playlists cluster together, suggesting a network connection between depression songs. Emotional reflection and religious or biblical language are dominant topics in depression playlist communities, which are different from non-depression playlist communities. This explorative analysis not only heightens our understanding of the nuanced relationships between music preferences and mental health but also illuminates pathways for cultivating sensitive, adaptive interventions and technologies in mental health discourse and practice.

### 4.2 Implications

For RQ1, our work showed that several musical patterns differ between depression and non-depression users. Our result is consistent with previous research on music and psychology, showing that depression patient tend to choose music with higher energy [29]. Moreover, previous research indicated that higher energy decrease the perception of relaxing for songs [5]. However, as mental health research suggests that relaxation can significantly reduce the level of depression [30], it raises an intriguing paradox: while individuals with depression are drawn to high-energy music, which might be less perceived as relaxing, they may unintentionally be selecting music that doesn't optimize their psychological well-being, given the documented benefits of relaxation for mitigating depressive symptoms. Moreover, our results found that songs shared by depression individuals have lower popularity. Previous research on depression also showed that preference for subculture elements

(such as anime, goths, punks) and less popular interests are positively associated with depressive symptoms [8, 11, 39]. This preference for less popular songs might stem from a variety of factors: the depressive mood might be reinforced by subcultures [13], and people with depression feel disconnected with popular culture [39]. Additionally, reduced preference for speechiness is coherent with the genres and lyrics analysis showing that songs shared by depression users are less likely to be rap and contain explicit and rap languages. This illustrates a consistent connection between the analysis from different aspects in our work.

For RQ2, our work qualitatively showed that the artists and genres commonly shared by depression and non-depression individuals are different. The result shared common ground with previous work indicating that soul music listening(hip hop and R&B) is a predictor of lower depression levels [41]. Hip-hop and R&B, often associated with expressions of resilience, empowerment, and an outlet for navigating through socio-emotional challenges, might not resonate with depression individuals' emotional and psychological state.

For RQ3, our results demonstrated that the topic and language use in depression and non-depression groups might be different. First, depression-prone individuals are more likely to share songs with topics in everyday life and emotional recollection. This is coherent with psychology research indicating that those with depression report more rumination than the control group [45]. However, rumination can lead to negative consequences in mood: rumination exacerbates and prolongs depressed mood [44], and rumination makes it harder for patients to recover from psychological disorders [42, 51]. We hypothesize that listening to songs with emotional recollection topic may induce ruminative thoughts, which, in turn, exacerbate depression. Moreover, the LIWC result indicated that depressed individuals prefer lyrics with short sentences. This is probably because depression decreases the cognitive load [43]. Therefore, short sentences are easier to be perceived and liked by depressed individuals.

Finally, for RQ4, we found that most of the communities display polarized with respect to either depression or non-depression playlists, which implies that the depression playlists may share similar music preference and cluster together. This may lead to the echo chamber effect, making depressed individuals being suggested by more depression songs, and further exacerbate and prolong depressed mood. Moreover, consistent with our topic modeling result, our network analysis also showed that the dominant topics of depression and non-depression communities are different. The

dominance of religious and biblical languages in depression community is coherent with both LDA result and previous research showing that depressed individuals engaged more with expression of religious involvement on social media [14]. In summary, our work validate that the music preference is different between depression and non-depression group from both single song and playlist level. Utilizing both LDA and community detection serves to corroborate our findings and enhance the validity of the result.

We discern crucial implications across two pivotal dimensions. Firstly, the characteristic features of individuals with depression in relation to their music preferences can be observed. On a social level, this group demonstrates a propensity for songs that may not be widely recognized or popular, possibly reflecting a diminished sense of belongingness and reduced proclivity for social engagement. This is also emphasized by their inclination away from soul music, often imbued with themes of resilience, empowerment, and socialization. From a cognitive perspective, a preference for shorter sentences may indicate a desire for easily understood content, potentially deriving from a lower cognitive load. Secondly, we delve into the consequences engendered by these specific musical preferences. Individuals with depression tend to gravitate towards high-energy music, paradoxically offering less relaxation, and exhibit an avoidance of genres articulating resilience. Moreover, a preference for songs with emotional reflection topics may inadvertently foster rumination, further entrenching depressive states. This is not only evident in the lyric analysis in the network analysis. Incorporating all aforementioned observations, the musical preferences of depression individuals inadvertently forge a self-reinforcing echo chamber, where the music perpetuates and potentially exacerbates prevailing depressive moods and emotions. This echo chamber effect is manifested not merely through individual listening choices but is further perpetuated by the recommendation algorithms that may suggest additional depression playlists to listeners. Thus, recommendation algorithms, influenced by these patterns, may unwittingly suggest additional depression-congruent content to listeners, thereby sustaining and potentially intensifying the resonance with depressive moods and emotions through continued exposure to thematically similar music.

Based on the result, we also suggest that music application develop algorithms or frameworks to gauge the emotional state of listeners, by analyzing their music listening patterns over time, thus providing an unintrusive way to monitor emotional well-being. Music application can inquire the mental health condition of their users and recommend songs or playlists with less depressive features, and explore the efficacy of such tailored musical interventions and whether guiding listening habits can serve as a viable strategy in emotional self-regulation or mental health management. In addition, the answer to the question "How can recommendation algorithms avoid negative reinforcement and echo chamber effect?", needs to be explored. These ideas may help developing holistic mental health supporting systems.

### 4.3 Limitation and Future Work

We acknowledge our study has several limitations, some of which point to intriguing future research directions. First, this study focused on the overall profile and preferences of depressed and non-

depressed users, while a potential future work can be eliminating the overlapping features and analyzing the distinctive unique features for each of the users classes. Second, since we do not have the playlist information from depression/non-depression users, we cannot analyze the actual music preference connection between depressed individuals in the network analysis. Instead, we have to analyze in an indirect way that regards playlists with a majority of songs shared by depressed individuals as depression playlists. A potential future direction could focus on collecting social network data with depression playlists to further study this question.

## REFERENCES

- [1] [n. d.]. Spotify Million Playlist Dataset Challenge. <https://www.aicrowd.com/challenges/spotify-million-playlist-dataset-challenge>. Accessed: 2023-09-27.
- [2] [n. d.]. twitter-depression-tweets-and-musics. <https://www.kaggle.com/datasets/rmmartin/twitter-depression-tweets-and-musics>. Accessed: 2023-09-27.
- [3] 2023. Web API. <https://developer.spotify.com/documentation/web-api>. Accessed: 2023-09-27.
- [4] Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive self-disclosures, responses, and social support on Instagram: the case of depression. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 1485–1500.
- [5] Margarida Baltazar and Daniel Västfjäll. 2020. Songs perceived as relaxing: Musical features, lyrics, and contributing mechanisms. In *International Conference: Psychology and Music—Interdisciplinary Encounters*. Faculty of Music, University of Arts in Belgrade.
- [6] Michael D Barone, Jotthi Bansal, and Matthew H Woolhouse. 2017. Acoustic features influence musical choices across multiple genres. *Frontiers in psychology* 8 (2017), 931.
- [7] Krishna C Bathina, Marijn Ten Thij, Lorenzo Lorenzo-Luaces, Lauren A Rutter, and Johan Bollen. 2021. Individuals with depression express more distorted thinking on social media. *Nature human behaviour* 5, 4 (2021), 458–466.
- [8] Nejra Bešić and Margaret Kerr. 2009. Punks, Goths, and Other Eye-Catching Peer Crowds: Do They Fulfill a Function for Shy Youths? *Journal of Research on Adolescence* 19, 1 (2009), 113–121.
- [9] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
- [10] Leonardo Bonetti and Marco Costa. 2016. Intelligence and musical mode preference. *Empirical Studies of the Arts* 34, 2 (2016), 160–176.
- [11] Lucy Bowes, Rebecca Carnegie, Rebecca Pearson, Becky Mars, Lucy Biddle, Barbara Maughan, Glyn Lewis, Charles Fernyhough, and Jon Heron. 2015. Risk of depression and self-harm in teenagers identifying with goth subculture: a longitudinal cohort study. *The Lancet Psychiatry* 2, 9 (2015), 793–800.
- [12] R. L. Boyd, A. Ashokkumar, S. Seraj, and J. W. Pennebaker. 2022. *The development and psychometric properties of LIWC-22*. Technical Report. University of Texas at Austin, Austin, TX. <https://www.liwc.app>
- [13] Rebecca C Brown and Paul L Plener. 2017. Non-suicidal self-injury in adolescence. *Current psychiatry reports* 19 (2017), 1–8.
- [14] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th annual ACM web science conference*, 47–56.
- [15] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, Vol. 7. 128–137.
- [16] Marc JMH Delsing, Tom FM Ter Bogt, Rutger CME Engels, and Wim HJ Meeus. 2008. Adolescents' music preferences and personality characteristics. *European journal of Personality: Published for the European Association of Personality Psychology* 22, 2 (2008), 109–130.
- [17] Tuomas Eerola and Jonna K Vuoskoski. 2011. A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music* 39, 1 (2011), 18–49.
- [18] Jaakko Erkkilä, Marko Punakanen, Jörg Fachner, Esa Ala-Ruona, Inga Pöntiö, Mari Tervaniemi, Mauno Vanhala, and Christian Gold. 2011. Individual music therapy for depression: randomised controlled trial. *The British journal of psychiatry* 199, 2 (2011), 132–139.
- [19] SAGS Evans-Lacko, Sergio Aguilar-Gaxiola, Ali Al-Hamzawi, Jordi Alonso, Corina Benjet, Ronny Bruffaerts, WT Chiu, Silvia Florescu, Giovanni de Girolamo, Oye Gureje, et al. 2018. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. *Psychological medicine* 48, 9 (2018), 1560–1571.
- [20] Bruce Ferwerda, Emily Yang, Markus Schedl, and Marko Tkalcic. 2015. Personality traits predict music taxonomy preferences. In *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems*, 2241–2246.

- [21] Maya B Flannery and Matthew H Woolhouse. 2021. Musical preference: Role of personality and music-related acoustic features. *Music & Science* 4 (2021), 20592043211014014.
- [22] Paula Frieiro Padín, Rubén González Rodríguez, María Del Carmen Verde Diego, Raquel Vázquez Pérez, et al. 2021. Social media and eating disorder psychopathology: A systematic review. *Cyberpsychology Journal of Psychosocial Research on Cyberspace* (2021).
- [23] Sandra Garrido and Emery Schubert. 2015. Music and people with tendencies to depression. *Music Perception: An Interdisciplinary Journal* 32, 4 (2015), 313–321.
- [24] Ian H Gotlib and Constance L Hammen. 2008. *Handbook of depression*. Guilford Press.
- [25] Alinka E Greasley and Alexandra M Lamont. 2006. Music preference in adulthood: Why do we like the music we do. In *Proceedings of the 9th international conference on music perception and cognition*. University of Bologna Bologna, Italy, 960–966.
- [26] David M Greenberg, Michal Kosinski, David J Stillwell, Brian L Monteiro, Daniel J Levitin, and Peter J Rentfrow. 2016. The song is you: Preferences for musical attribute dimensions reflect personality. *Social Psychological and Personality Science* 7, 6 (2016), 597–605.
- [27] Sharath Chandra Guntuku, J Russell Ramsay, Raina M Merchant, and Lyle H Ungar. 2019. Language of ADHD in adults on social media. *Journal of attention disorders* 23, 12 (2019), 1475–1485.
- [28] Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar, and Johannes C Eichstaedt. 2017. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences* 18 (2017), 43–49.
- [29] Claire Howlin and Brendan Rooney. 2021. Patients choose music with high energy, danceability, and lyrics in analgesic music listening interventions. *Psychology of Music* 49, 4 (2021), 931–944.
- [30] Yong Jia, Xiaowei Wang, and Yuanjuan Cheng. 2020. Relaxation therapy for depression: an updated meta-analysis. *The Journal of Nervous and Mental Disease* 208, 4 (2020), 319–328.
- [31] Benjamin K Johnson and Giulia Ranzini. 2018. Click here to look clever: Self-presentation via selective sharing of music and film on social media. *Computers in Human Behavior* 82 (2018), 148–158.
- [32] Steve Jones. 2000. Music and the Internet. *Popular Music* 19, 2 (2000), 217–230.
- [33] Anahid Kassabian. 2013. *Ubiquitous listening: Affect, attention, and distributed subjectivity*. Univ of California Press.
- [34] Carol L Krumhansl. 2002. Music: A link between cognition and emotion. *Current directions in psychological science* 11, 2 (2002), 45–50.
- [35] Dongwon Lee, Jaemie Yejean Park, Junha Kim, Jaejeung Kim, and Junghoon Moon. 2011. Understanding music sharing behaviour on social network services. *Online Information Review* 35, 5 (2011), 716–733.
- [36] David Lester and Melissa Whipple. 1996. Music preference, depression, suicidal preoccupation, and personality: Comment on Stack and Gundlach's papers. *Suicide and Life-Threatening Behavior* 26, 1 (1996), 68–70.
- [37] Daniel Leubner and Thilo Hinterberger. 2017. Reviewing the effectiveness of music interventions in treating depression. *Frontiers in psychology* 8 (2017), 1109.
- [38] Tony Liu, Jonah Meyerhoff, Johannes C Eichstaedt, Chris J Karr, Susan M Kaiser, Konrad P Kording, David C Mohr, and Lyle H Ungar. 2022. The relationship between text message sentiment and self-reported depression. *Journal of affective disorders* 302 (2022), 7–14.
- [39] Yinghao Liu, Yingxu Liu, and Jiahao Wen. 2022. Does anime, idol culture bring depression? Structural analysis and deep learning on subcultural identity and various psychological outcomes. *Heliyon* 8, 9 (2022).
- [40] Anna Maratos, Christian Gold, Xu Wang, and Mike Crawford. 2008. Music therapy for depression. *Cochrane database of systematic reviews* 1 (2008).
- [41] Dave Miranda and Michel Claes. 2008. Personality traits, music preferences and depression in adolescence. *International journal of adolescence and youth* 14, 3 (2008), 277–298.
- [42] Rebecca Morrison and Rory C O'Connor. 2008. A systematic review of the relationship between rumination and suicidality. *Suicide and Life-Threatening Behavior* 38, 5 (2008), 523–538.
- [43] Stevan Nikolin, Yi Yin Tan, Andreas Schwaab, Adriano Moffa, Colleen K Loo, and Donel Martin. 2021. An investigation of working memory deficits in depression using the n-back task: A systematic review and meta-analysis. *Journal of Affective Disorders* 284 (2021), 1–8.
- [44] Susan Nolen-Hoeksema and Jannay Morrow. 1993. Effects of rumination and distraction on naturally occurring depressed mood. *Cognition & emotion* 7, 6 (1993), 561–570.
- [45] Bunmi O Olatunji, Kristin Naragon-Gainey, and Kate B Wolitzky-Taylor. 2013. Specificity of rumination in anxiety and depression: A multimodal meta-analysis. *Clinical Psychology: Science and Practice* 20, 3 (2013), 225.
- [46] Renato Panda, Hugo Redinho, Carolina Gonçalves, Ricardo Malheiro, and Rui Pedro Paiva. 2021. How does the spotify api compare to the music emotion recognition state-of-the-art? In *18th Sound and Music Computing Conference (SMC 2021)*. Axeia sas/SMC Network, 238–245.
- [47] Esteban A Ríssola, Mohammad Aliannejadi, and Fabio Crestani. 2022. Mental disorders on online social media through the lens of language and behaviour: Analysis and visualisation. *Information processing & management* 59, 3 (2022), 102890.
- [48] Jenefer Robinson. 2005. *Deeper than reason: Emotion and its role in literature, music, and art*. Oxford University Press.
- [49] Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports* 9, 1 (2019), 5233.
- [50] Raquel Campos Valverde. 2022. Online musicking for humanity: the role of imagined listening and the moral economies of music sharing on social media. *Popular Music* 41, 2 (2022), 194–215.
- [51] Edward R Watkins and Henrietta Roberts. 2020. Reflecting on rumination: Consequences, causes, mechanisms and treatment of rumination. *Behaviour Research and Therapy* 127 (2020), 103573.
- [52] Richard M Wenzlaff and Danielle E Bates. 1998. Unmasking a cognitive vulnerability to depression: how lapses in mental control reveal depressive thinking. *Journal of personality and social psychology* 75, 6 (1998), 1559.
- [53] WHO. 2023. Depressive disorder (depression). <https://www.who.int/news-room/fact-sheets/detail/depression>. Accessed on: 2023-09-24.
- [54] Kay Wilhelm, Inika Gillis, Emery Schubert, and Erin Louise Whittle. 2013. On a blue note: depressed peoples' reasons for listening to music. *Music and Medicine* (2013).
- [55] CA Woody, AJ Ferrari, DJ Siskind, HA Whiteford, and MG Harris. 2017. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. *Journal of affective disorders* 219 (2017), 86–92.
- [56] Ronghua Xu and Qingpeng Zhang. 2016. Understanding online health groups for depression: social network and linguistic perspectives. *Journal of medical Internet research* 18, 3 (2016), e63.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009