

Original Paper

Resilience in Web-Based Mental Health Communities: Building a Resilience Dictionary With Semiautomatic Text Analysis

Yong-Bin Kang¹, PhD; Anthony McCosker^{1,2}, PhD; Peter Kamstra², PhD; Jane Farmer², PhD

¹Australian Research Council (ARC) Centre of Excellence for Automated Decision-Making and Society (ADM+S), Swinburne University of Technology, Victoria, Australia

²Social Innovation Research Institute, Swinburne University of Technology, Victoria, Australia

Corresponding Author:

Yong-Bin Kang, PhD

Australian Research Council (ARC) Centre of Excellence for Automated Decision-Making and Society (ADM+S)

Swinburne University of Technology

John St

Hawthorn

Victoria, 3122

Australia

Phone: 61 3 9214 5904

Email: ykang@swin.edu.au

Abstract

Background: Resilience is an accepted strengths-based concept that responds to change, adversity, and crises. This concept underpins both personal and community-based preventive approaches to mental health issues and shapes digital interventions. Online mental health peer-support forums have played a prominent role in enhancing resilience by providing accessible places for sharing lived experiences of mental issues and finding support. However, little research has been conducted on whether and how resilience is realized, hindering service providers' ability to optimize resilience outcomes.

Objective: This study aimed to create a resilience dictionary that reflects the characteristics and realization of resilience within online mental health peer-support forums. The findings can be used to guide further analysis and improve resilience outcomes in mental health forums through targeted moderation and management.

Methods: A semiautomatic approach to creating a resilience dictionary was proposed using topic modeling and qualitative content analysis. We present a systematic 4-phase analysis pipeline that preprocesses raw forum posts, discovers core themes, conceptualizes resilience indicators, and generates a resilience dictionary. Our approach was applied to a mental health forum run by SANE (Schizophrenia: A National Emergency) Australia, with 70,179 forum posts between 2018 and 2020 by 2357 users being analyzed.

Results: The resilience dictionary and taxonomy developed in this study, reveal how resilience indicators (ie, "social capital," "belonging," "learning," "adaptive capacity," and "self-efficacy") are characterized by themes commonly discussed in the forums; each theme's top 10 most relevant descriptive terms and their synonyms; and the relatedness of resilience, reflecting a taxonomy of indicators that are more comprehensive (or compound) and more likely to facilitate the realization of others. The study showed that the resilience indicators "learning," "belonging," and "social capital" were more commonly realized, and "belonging" and "learning" served as foundations for "social capital" and "adaptive capacity" across the 2-year study period.

Conclusions: This study presents a resilience dictionary that improves our understanding of how aspects of resilience are realized in web-based mental health forums. The dictionary provides novel guidance on how to improve training to support and enhance automated systems for moderating mental health forum discussions.

(*JMIR Form Res* 2022;6(9):e39013) doi: [10.2196/39013](https://doi.org/10.2196/39013)

KEYWORDS

resilience dictionary; mental health; peer-support forum; topic modeling; text analysis; content moderation

Introduction

Background

Mental health is fundamental to our individual and collective capabilities of thinking and interacting with each other and enjoying life [1]. Mental health challenges have been increasing worldwide, with consequent negative impacts on social and economic prosperity [2-4]. A 2017 survey [5] found that approximately 1 in 7 people (equivalent to 1 billion people) worldwide have experienced mental illness. In Australia, government statistical reports show that approximately 20% of people (4.8 million) experience mental health issues such as anxiety-related conditions and depression [6]. From 2020 to 2021 during the COVID-19 pandemic, 3.4 million Australians aged 16 to 85 years (17% of that age group) sought help from a mental health professional [7]. With its lockdowns and social distancing requirements, the COVID-19 pandemic is associated with what some have called an “unprecedented mental health crisis” [8], accelerating the need for community and digital interventions [9].

Reaching out to others for support is a vital part of enhancing resilience [10]. On this basis, it has been shown that *online peer-support mental health forums* (eg, those run in Australia by the charities Beyond Blue [11], SANE [Schizophrenia: A National Emergency] Australia [12], and ReachOut [13]) play increasingly important roles in establishing social connections, sharing knowledge and experiences, and providing emotional support among people with lived experiences of mental illness [4,14]. Such forums complement publicly funded and private health services by enabling firsthand access to people with shared experiences, advice, and guidance. Within forums, specially trained staff, volunteers, and untrained peers provide a supportive and safe space to talk and be heard [15,16]. Evidence shows that the benefits include the following [17]: (1) building safe and trusting relationships, (2) ensuring values of mutuality and reciprocity, (3) promoting validation and application of experiential knowledge, (4) enabling peers to exercise leadership through peer support, and (5) empowering peers to discover and make use of their own strengths. As complements to clinical and telehealth services, there remains a need to better understand the strengths of peer-support forums and optimize their management and moderation.

Resilience is an accepted strengths-based concept rather than a deficit or harms-based concept. Generating resilience is a regularly suggested approach to addressing mental health issues in community settings and falls within a preventive model of health care [8,18,19]. Resilience is understood variously across several disciplines as an individual and collective capability to adapt in the face of adversity, trauma, tragedy, threats, or even significant sources of stress [20]. In response to challenges, resilient individuals and communities can draw on psychosocial resources—personal and collective—to cope and adapt [4]. As a result, an individual’s resilience can positively affect their mental health. However, more research is needed to understand the characteristics of resilience and how it is realized through interactions within dedicated web-based mental health support settings, such as peer-support forums. This research contributes

a method, as well as findings, that can help forum service providers demonstrate their outcomes and tailor their moderation practices to optimize resilience-building interactions.

Prior Work

Despite the potential utility of online mental health peer-support forums for building resilience, *whether and how resilience is realized* by participants has received little research attention. Where resilience or strengths-based resources are the focus of existing studies, approaches fall into two streams: (1) qualitative analysis using user surveys and (2) qualitative and quantitative analysis of web-based forum data using either manual or larger-scale artificial intelligence techniques. The first stream is centered on manually identifying aspects of resilience and their reciprocal associations with mental health from web-based user surveys during particular events (eg, the COVID-19 pandemic [9,19]) or from certain groups (eg, university students [18] or adolescents [21]). User survey data are less adept at showing how resilience is realized through forum interaction and discourse as it is cross-sectional at points in time and asks specific questions created by survey designers, which might not resonate with forum users’ perceptions and experiences.

The second research stream has evolved with recent advances in natural language processing (NLP) and machine learning. These methods have shown promise as a way of using forum content as data sets for exploring public health questions [22]. However, much of this work has been “risk” and “harms” focused rather than strengths focused. For example, a triaging system was used to assist web-based peer support by classifying forum messages into different risk levels based on how urgently moderators’ attention was needed [23]. A study [3] compared 2 web-based depression forums (Beyond Blue [11] and *r/depression* on Reddit [24]) using NLP techniques based on user sentiments and discussion topics. NLP techniques have also been used to identify which posts from web-based health forums (HealthBoards [25], Inspire [26], and HealthUnlocked [27]) are related to the COVID-19 pandemic based on conversations among mental health consumers. Other studies have applied NLP to detect sentiments and emotions (ie, fear, anger, sadness, and joy) [28-30] for depression diagnosis [31,32] and to understand grief processes [33] and stress [34]. As noted, these studies tend to target risks associated with mental ill health, leaving a gap in identifying strengths-based interactions in web-based data sets and how web-based services might help generate positive health outcomes.

Aims of This Study

As mental health services seek to optimize and expand digital support interventions, new methods of analysis and monitoring of forum activities are needed. In this study, we harnessed NLP techniques to reveal indicators of resilience and how resilience is realized within online peer-support forums. This work provides new insights into how resilience is realized in forums by building a *resilience dictionary* that reveals the resilience themes discussed across 2 years of forum activity.

More specifically, the resilience dictionary we generated shows (1) the types of themes that characterize resilience and (2) the relationships between different resilience indicators with respect

to their realization. These insights reveal the broader (more abstract or common) and narrower (more specific) nature of specific resilience indicators. The resilience dictionary can be used by forum service providers and researchers to show that resilience is enabled by forums and how this occurs. It can help monitor changes in resilience realized by forum users over time and understand what prompts resilience changes, both within the forum and stimulated by external events. By establishing a method for quantifying resilience in unstructured text data, the resilience dictionary can help forum managers and creators to think more strategically about how to design the forums and maintain and moderate them more effectively. Revealing the characteristics of resilience and the extent to which it is realized over time can also help nonprofit organizations attract funding to sustain and further develop these forums.

To achieve our aims, we investigated the following research questions (RQs):

- RQ1: How can we characterize resilience indicators using topics (ie, major themes of posts) and their descriptive terms?
- RQ2: How are resilience indicators realized over time, and which are more dominant than the others?
- RQ3: How can we create a resilience dictionary that reflects the characteristics of each resilience indicator and the relatedness between the resilience indicators?

To address these RQs, we conducted a semiautomatic approach using methods that incorporate *topic modeling* (a type of statistical modeling in NLP) and exploit *human knowledge*.

To address RQ1, we explored optimal methods for discovering major thematic concepts, or *topics*, in forum data using topic modeling. Subsequently, using qualitative methods, we mapped these topics to the resilience indicators. The topics and their descriptive terms formed the controlled vocabulary of the resilience dictionary.

To investigate RQ2, we observed the *resilience prevalence* by analyzing the proportion of input forum data during the specified periods. We then analyzed the prevalence of each resilience indicator over time.

To explore RQ3, we created a resilience dictionary using the results of topics mapped to each resilience indicator. To identify the relatedness between resilience indicators in terms of their realization, we examined the co-occurrence patterns of resilience indicators from the input forums to automatically build a taxonomy of resilience indicators (ie, *resilience taxonomy*). This resilience taxonomy sheds light on whether resilience indicators are dependent on or independent of other resilience indicators.

For this work, we used data produced by SANE Australia's 2 forums. SANE forums represent one of the largest web-based mental health peer-support communities in Australia, where people aged ≥ 18 years with mental health issues can register and engage in support, training, and education. Users can interact with other people (ie, peers) experiencing similar mental health challenges through the forums.

To our knowledge, this study is the first to explore the construction of a resilience dictionary by examining relationships between topics and resilience indicators and analyzing the relatedness between resilience indicators from web-based mental health forums using NLP and human knowledge. The significance of this study lies in the following aspects. First, we present how strengths-based resilience develops in web-based mental health forums based on a wider spectrum of web-based mental health discussions. Prior studies have attempted to accomplish this; however, they were small in scale and limited by specific events (eg, the COVID-19 pandemic) [8,9,14] or by certain age groups [2,18,21]. Second, the resilience dictionary can provide evidence-based information that can help design and maintain better mental health support services, as well as improve responses to people seeking mental health support. In particular, this benefit can be valuable for nonprofit mental health organizations that often have difficulty securing and allocating resources to improve their services. Finally, this work shows how mental health services can be made more effective by using support forum data, the value of which is becoming increasingly recognized [22,30].

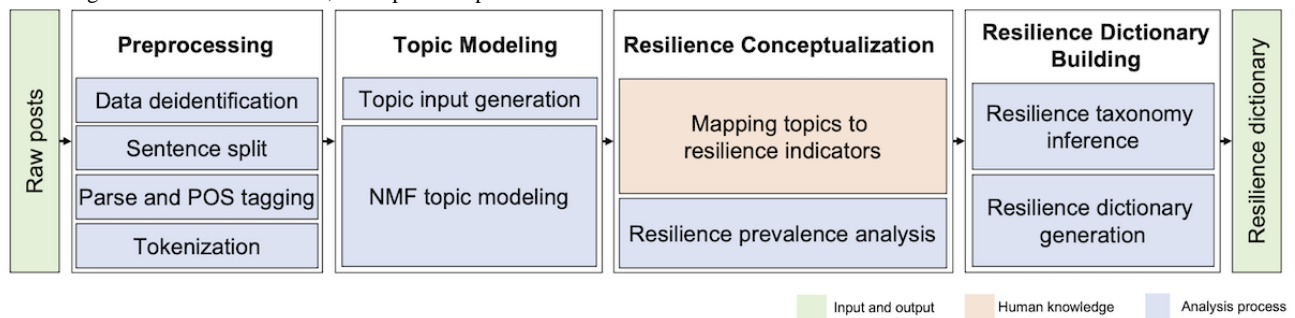
Methods

Overall Design

Overview

The 4 phases of our semiautomatic approach to creating a resilience dictionary are shown in [Figure 1](#). Forum posts were the *first* preprocess to ensure deidentification and perform tokenization to generate meaningful terms. *Second*, topic modeling was used to discover the key topics (ie, themes) discussed in the posts. *Third*, the topics were mapped to a framework of resilience indicators using thematic coding processes. *Fourth*, a resilience taxonomy was established based on the co-occurrence analysis of resilience indicators from the posts. A resilience dictionary was then generated by integrating the resilience taxonomy with all the outcomes of the previous 4 phases. In the following sections, we present a detailed description of the phases after introducing the data set and the nature of the resilience indicators used in the study.

Figure 1. The 4-phase analysis pipeline that preprocesses forum posts, finds topics, conceptualizes resilience indicators, and builds a resilience dictionary. NMF: nonnegative matrix factorization; POS: parts of speech.



SANE Forum Data

SANE Australia is a national mental health charity that supports Australians affected by mental illness by providing peer-support services, training, and counseling. SANE data were collected from two web-based forums: (1) the “Lived Experience Forum” and (2) the “Friend, Family and Carers Forum.” The former is for people who live with complex mental health issues, whereas the latter is for their supporters and carers. SANE forums are moderated by mental health practitioners and aim to offer a safe and anonymous space on the web for discussing mental health and related issues [35]. Forums provide participants with social connections, allow the sharing of feelings and emotions, and provide a source of information generated through conversations. Forum discussions are intended to help consumers affected by mental illness explore positive pathways to address issues affecting their lives. According to the SANE annual report for 2020, there were 21,041 registered forum members, and 151,137 individual Australians used the forums in 2020 [35].

We obtained a national sample of posts from the SANE forums from 2018 Q3 (starting July 2018) to 2020 Q4 (ending December 2020). In total, 70,179 posts by 2357 users were obtained and used in this study in collaboration with the SANE forum managers.

Resilience Indicators

Drawing on the resilience framework proposed by Berkes and Ross [36], which combines psychological and community development approaches, we identified common components of personal and community resilience. Berkes and Ross [36] devised a conceptual framework of resilience pertaining to people’s lives in communities, allowing the concept to account

for the wider ecosystem of factors affecting an individual’s mental health and well-being. Although there is a range of approaches to resilience, this framework is useful for understanding how resilience is realized by members of a forum, given their location in both physical on-the-ground communities and the web-based peer-based community of the forums.

The adopted resilience approach is premised on the understanding that people and communities that are resilient require access to a set of personal, interpersonal, or community resources. On the basis of research evidence, these involve *learning* or access to new knowledge, information, and skills, leading to an increased capability of dealing with change [37]; *social capital* or access to networks of people and the support, trust, and integration they can foster; a sense of *belonging*, including belonging to groups and places, acceptance as part of a group, and the processes that lead to identity formation [38]; *adaptive capacity* or using resources to enable adaptation and positive behavior change [39]; and *self-efficacy* or being able to self-organize or work toward feeling in a controlled state [36]. The definitions of these indicators and how they are conceptually realized in the context of mental health are shown in Table 1. In this study, the generation of the resilience dictionary was focused on and built from these 5 indicators.

Theoretically, we expect to see some variation in the way each indicator of resilience manifests in the web-based discussion context (in comparison with other contexts) and differences in their prevalence [36]. This means that some indicators should act as generators or resources, and others act as attributes or characteristics of resilience capacity, such as adaptive capacity or instances of self-efficacy. These dependencies were tested through RQ3, with the application of the taxonomic analysis described in phase 4.

Table 1. Resilience indicators, their definitions, and their conceptual realization in web-based forums.

Resilience indicators	Definition	Realization in community forums
Social capital [38]	Access to social networks and the support, trust, and social integration they can foster	<ul style="list-style-type: none"> • Formation of social connections that individuals could draw on for confidence and support • Expressing a willingness to share stories of similar experiences with mental health issues
Belonging [38]	Belonging to people and places, acceptance as part of a group, and the processes that lead to identifying formation	<ul style="list-style-type: none"> • Introducing themselves to others • Sharing stories and lived experiences with those experiencing similar challenges and participating at their own pace
Learning [37]	Access to knowledge, information, and skill development	<ul style="list-style-type: none"> • Sharing applied and experiential knowledge with others • Offering advice on how to practice good mental health and strategies on how to navigate mental health services
Adaptive capacity [39]	Use of resources that enable adaptation and positive behavior change	<ul style="list-style-type: none"> • Self-reporting their own activity or behavior changes when responding to advice or information on the forum
Self-efficacy [36]	Ability to self-organize or work toward feeling in a controlled state	<ul style="list-style-type: none"> • Communicating feelings of control in the aspects of their life or condition • Supporting others by keeping them accountable for their changed behaviors or positive lifestyle choices that made them feel in control

Ethics Approval

The university ethics committee approved this study (R/2019/033). In addition, we adhered to SANE's ethics and data governance policies throughout, which included establishing a data-sharing agreement, anonymizing posts to protect forum users' identities, and applying a data security protocol.

Phase 1: Preprocessing of Forum Posts

In the preprocessing phase, the goal was to break up each post into smaller basic units of meaning. User identifiers (IDs or names) were removed to deidentify the posts. Posts were then split into sentences to perform part-of-speech tagging and remove stop words (words such as *is*, *are*, and *the*, which do not carry useful information). The remaining terms were converted into their lemmatized forms to group the inflected forms of words together. Posts that were too short to add value to topic modeling (<5 terms) were removed. Finally, the remaining terms were used as the input to the next "topic modelling" phase.

Phase 2: Topic Modeling

Overview

Topic modeling has gained attention because of its capacity to discover latent thematic topics in a large corpus of text documents [40,41]. A topic is a specific, recognizable theme defined by a cohesive set of terms representing the characteristics of that theme. Nonnegative matrix factorization (NMF) [42] was chosen as a method because of its effectiveness in discovering topics in short text documents and its ability to discover both broad and specific topics. As a method of analyzing forum posts, NMF can help identify both broad themes discussed often over time and more specific themes that might relate to periodic events [43].

A key idea of NMF is to decompose a term-document matrix n by m A (n is the number of forum posts, and m is the number of terms) into 2 nonnegative submatrices, n by k W (k is the number of topics) and k by m H , such that A is approximated by the multiplication of W and H , denoted as $A \approx W \times H$. In A , the weight of a term in a document is measured by a well-defined method, the term frequency-inverse document frequency weighting scheme [44], via term frequency and rarity. The matrix W represents the document membership weights over the topics. Each row denotes a document, and each column corresponds to a topic. Sorting the values of a column (topic) provides the ranking of the most relevant documents for the column. The matrix H contains the term weights relative to each of the topics. A row defines a topic. Sorting the values on each row provides the ranking of the most relevant terms (descriptors) of each topic.

Topic Input Generation

After preprocessing, each post was represented by a list of terms. Before applying topic modeling, we performed 2 steps to generate topic inputs. In the first step, the posts were split into disjoint collections based on the time duration of a "quarter" of a year. This collection period was the best fit for identifying granular topics (ie, neither too broad nor too specific). This generated 10 post collections from 2018 Q3 to 2020 Q4. The second step involved building a term-document matrix A using the term frequency-inverse document frequency weighting scheme from each collection.

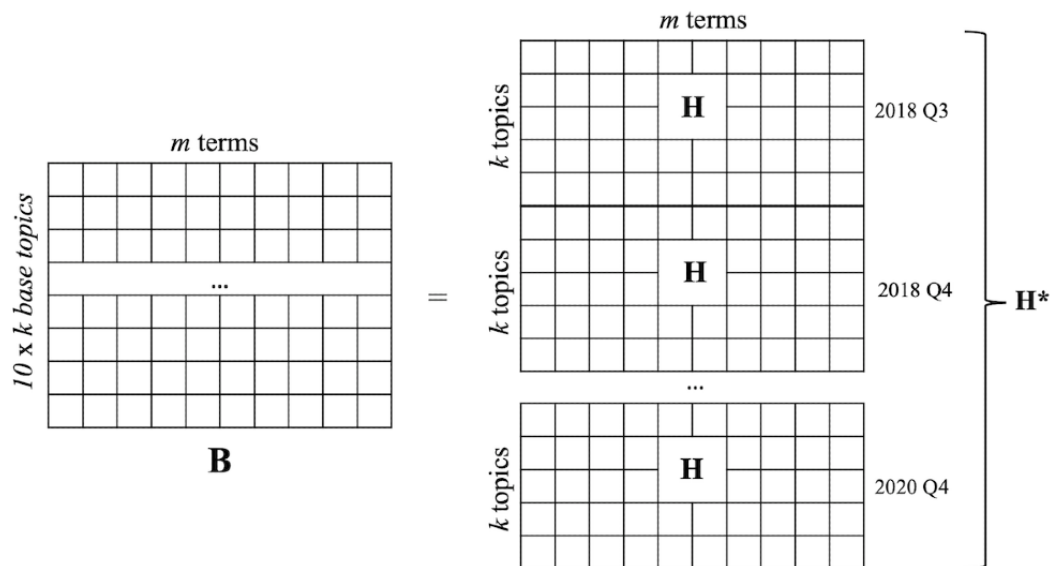
Two-Layered NMF Topic Modeling

The 2-layered NMF approach [45] was applied to term-document matrices (denoted by A^*) to discover both broader (longer-lived) and more specific (shorter-lived) topics. This process extracts broader topics observed across the entire time frame (2018 Q3 to 2020 Q4), as well as specific topics in each quarterly period, generating 2 topic layers. In the first layer, we find k topics, called *base topics*, from each term-document

matrix A of A^* using NMF, where k is a user-specified parameter. The output of this layer is a set of 2 matrices, W and H , from all disjoint post collections, denoted by W^* and H^* , respectively. In the second layer, we identified another type of topic, called *ensemble topics*, by analyzing similarities and variations among all the base topics generated from 2018 Q3 to 2020 Q4. A base topic-term matrix B was created by stacking each matrix H from H^* (Figure 2). Each H comprises k topics

with their constituent terms. In B , each row corresponds to a base topic, and each column is a term from the original posts. Each entry in B shows the weight (or importance) of the association of a term with a base topic, where the weight was inferred by the first-layer NMF. As shown in Figure 2, the dimension of B is $10 \times k$ by m , as H^* , which comprises 10 different matrices of H (from 2018 Q3 to 2020 Q4) and m terms.

Figure 2. Conceptual illustration of constructing the matrix B by stacking up the matrices of H^* generated by the first-layer nonnegative matrix factorization.



From B , we applied NMF to produce k' ensemble topics, where k' is a user-specified parameter, thus producing $B \approx W' \times H'$, where (1) W' denotes the base topic membership weights over the k' ensemble topics, and (2) the H' matrix shows the weights of terms for each ensemble topic. To produce the membership weights of the original posts over an ensemble topic, we calculated the following expression:

$$D \approx C \times W'^T \text{ (equation 1)}$$

Here, C is the document-base topic matrix made by stacking each matrix W of W^* in the same way as B was created from the matrices of H^* . D is the document-ensemble topic matrix and W'^T is the transpose of W' . Each row in D is an original post, and each column represents an ensemble topic.

Phase 3: Resilience Conceptualization

Overview

In this section, we present how to conceptualize each resilience indicator using its relevant topics. In this context, conceptualization refers to the process of specifying the characteristics of resilience indicators based on their relevant discussion themes (ie, topics) from qualitative forum posts. First, we present how to characterize each indicator using the discovered topics. Second, we present a method to determine whether resilience indicators are time varying. This can provide additional insights into their dynamic characteristics over time.

Mapping Topics to Resilience Indicators

Here, our goal was to connect topics with relevant resilience indicators (ie, in relation to our predetermined resilience

framework discussed previously). For this, we conducted qualitative content analysis, where knowledge of what happens on the forum and its aims is used to assess the meanings of topics and link each topic to relevant indicators. As qualitative data, we analyzed the meaning of the top N (N is a user-specified parameter) descriptive terms for each topic. As complementary data, we also examined the posts most relevant to each topic to analyze the meaning of the topic more rigorously in relation to the indicators of resilience. This content analysis was deductively performed using the framework of predefined resilience indicators. Note that a topic can be mapped to >1 resilience indicator depending on the thematic analysis. We followed a recent work showing the benefits of combining topic modeling with qualitative methods, particularly in the interpretation and contextualization phases of the analysis [46]. Here, our focus was on gaining an understanding of the nature of topics and their associations with resilience indicators based on domain knowledge of the context and posts that contain them most frequently. Topics cannot be meaningfully interpreted based only on the top N words as topics are not independent of the posts in which they appear.

The two types of qualitative data generated and analyzed were as follows.

First, the top N terms that best describe each ensemble topic (henceforth simply, “topic”) were identified from the topic-term matrix H' generated from the matrix B (recall $B \approx W' \times H'$ presented in phase 2; Figure 2). Note that each row in H' indicates the weights of all terms with respect to a topic. We generated these top N terms by ranking the weights. Examining

another resilience indicator y if the posts annotated with y are a subset of the posts annotated with x . On the basis of this scheme, we can find taxonomic relations between resilience indicators.

Resilience Dictionary Generation

By aggregating the findings to this point, we generated a resilience dictionary that represents (1) how resilience indicators

are characterized by particular topics, (2) how each topic is represented by specific descriptive terms and what terms are similar, and (3) what semantic relationships exist between resilience indicators. The resilience dictionary can improve the understanding of the semantic coverage of the meaning and realization of each resilience indicator. The dictionary comprises the features described in [Textbox 1](#).

Textbox 1. Features of the resilience dictionary.

Resilience dictionary features	
•	<i>Resilience indicator</i> indicates a resilience indicator.
•	<i>Parent</i> indicates the direct parent resilience indicator for a given indicator identified in the resilience taxonomy.
•	<i>Topic</i> provides a list of topics relevant to each resilience indicator determined using human knowledge.
•	<i>Topic weight</i> indicates the average weight of each topic over time. To calculate this weight, we first found the topic most relevant to each post by identifying the highest membership weight of the post over topics from the document-ensemble topic matrix D . In a sense, this weight is also seen as the weight of the topic (the most significant topic) in the post. Second, we calculated the average weight of the most significant topics across all the posts. Thus, this feature represents the relative importance of each topic over all the topics.
•	<i>Topic word</i> denotes the top 10 most descriptive terms of each topic.
•	<i>Similar terms</i> characterize the synonyms of a topic term, aiming to increase the semantics of a topic term. As there is no existing synonym dictionary for web-based mental health communities, our approach exploits a machine learning technique called <i>word embedding</i> . Word embedding is used to identify synonyms of each topic term based on their co-occurrence with other terms in input forum posts. Note that word embedding tends to indicate similar words by identifying the nearest words that appear together in similar contexts. Thus, synonyms captured by word embedding may not be “pure” synonyms depending on the input context (eg, if 2 terms “good” and “bad” frequently co-occur together in similar contexts, these can be identified as synonyms). We used a model named word2vec [50] as a word-embedding model, which has demonstrated many advantages for analyzing the semantic analysis of words.

Results

Phase 1: Preprocessing of Mental Health Forum Posts

From the original 70,179 posts posted by 2357 users, [Figure 4](#) shows the number of posts by time quarter. As the input for topic modeling, we used 69.56% (48,819/70,179) of posts. Note that we removed 30.44% (21,360/70,179) of posts with a length of <5 words (ie, too short).

[Figure 5](#) shows the top 50 terms by word frequency in the post data. Most comprise a mixture of nouns (eg, *day*, *help*, *need*, *way*, and *year*), verbs (eg, *help*, *need*, *feel*, *know*, *think*, *want*, and *sorry*), and adjectives (eg, *good*, *hard*, *great*, and *little*), whereas adverbs were relatively less used (eg, *really*). In total, from the 48,819 posts, we extracted 14,938 terms after the preprocessing phase. The minimum, average, and maximum term counts of the posts were 6, 30, and 1182, respectively.

Figure 4. Summary of Schizophrenia: A National Emergency (SANE) forum posts in the sample 2018-Q3 to 2020-Q4.

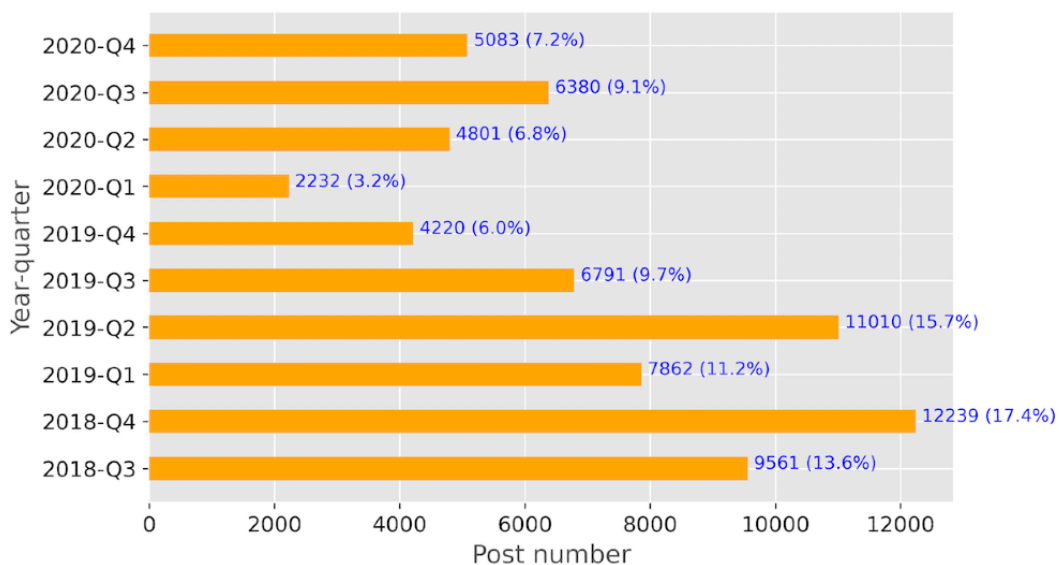
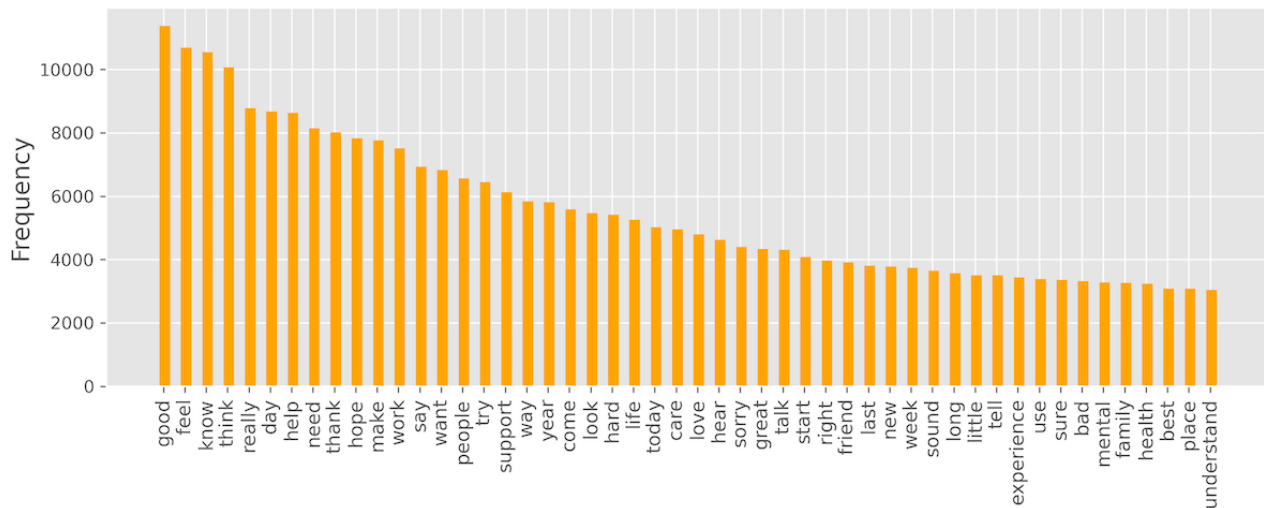


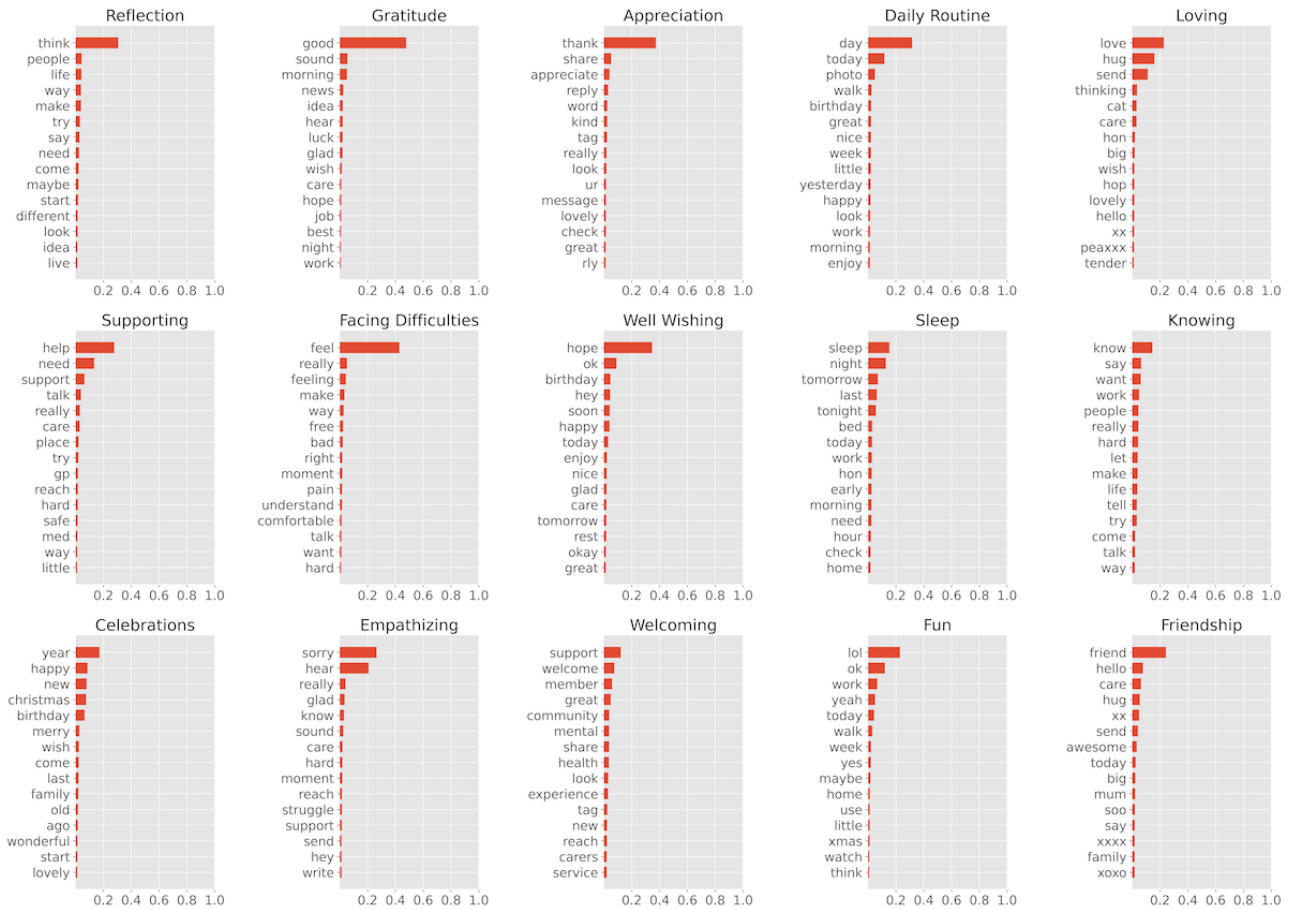
Figure 5. The top 50 terms by word frequency observed from the 48,819 posts.

Phase 2: Topic Modeling

The 48,819 posts were split into 10 disjoint post collections based on quarters 2018 Q3 to 2020 Q4. From each collection, we discovered topics using NMF. To generate topics, k and k' must be provided, where k and k' determine the number of base and ensemble topics to be generated, respectively. There is no universal method of determining such numbers [43]. In our study, we used a fixed value of 10 as k , assuming that we equally discovered the 10 main topics being discussed from each collection. Initial experiments found that $k=10$ (range was $k=5-15$) produced an informative topic set that was neither too general nor too specific. To discover the k' ensemble topics, a more complex method was applied by varying k' from 10 to 20 in increments of 1 to choose an optimal number. As we did not know how many base topics were similar or different, we measured similarities and variances of the base topics over the entire timeline using a widely used metric, *topic coherence* [51].

This method measures the degree of similarity between the top N terms for each topic. From the perspective of topic coherence, a better-interpretable topic yields a higher average score for the topic coherence scores of all generated topics. To measure topic coherence, we used word embedding [43] to measure the similarity between the top N terms in a topic. This similarity was estimated using the cosine similarity between word-embedded vectors. For the value of N , we used 10, following the suggestion in the study by O'Callaghan et al [43]. We generated a word-embedding model from the entire post data using word2vec [50]. Finally, using topic coherence, we selected 15 as the optimal value for k' . Figure 6 shows the 15 ensemble topics generated. For each topic, the top 15 descriptive terms were ranked vertically (the first row is the most descriptive term) based on their weights (importance). The weight of each term in a topic is visualized by a "red bar," where the sum of the weights of all terms in a topic is normalized to 1.

Figure 6. The 15 generated ensemble topics, where each topic is represented by its top 15 descriptive terms.



Phase 3: Resilience Conceptualization

Results of Mapping Topics to Resilience Indicators

A qualitative thematic analysis following a previous work [52] was conducted by 4 researchers with expertise in qualitative mental health research (second, third, and fourth authors) to interpret the meaning of the topics and assign descriptive topic labels to them. These labels, as shown in Figure 6, describe each topic by finding the most suitable meaning conveyed by the top 15 words in the identified topics and the top 20 posts associated with that topic. First, the researchers familiarized themselves with the topic modeling results, independently noting the initial ideas that were used to provide a semantically representative label of topic words and posts. Subsequently, topic labels were compared and refined to achieve satisfactory interresearcher agreement. The final agreed topic labels were then attributed to ≥ 1 of the 5 resilience indicators by each researcher through a deductive process based on the established resilience theory. Finally, these groupings were reviewed by the team and refined for consistency (Table 2).

The identified topics reflect the forum’s aim to provide ongoing, accessible, peer-to-peer support for people experiencing mental health conditions. Although specific instances of mental ill health, such as depression and anxiety, were often discussed, they were not the focus of the identified topics; that is, although

some topics expressed mental health problems directly, such as “sleep” and “facing difficulties,” daily challenges and indicators of social connection and support were more commonly identified topics. Topics also have an important temporal dimension, which is an advantage of our methodology, which considers topics over time. The 2-year sample of posts and topic modeling based on quarters meant that annual events or celebrations such as Christmas, New Year’s Eve, and birthdays were commonly referenced, showing the importance of social connection over time as a feature of the forum. Most topics expressed the establishment and maintenance of interpersonal social connection: “welcoming,” “empathising,” “well wishing,” “supporting,” “loving,” “friendship,” and “appreciation.” A few topics were more self-directed: “gratitude,” “reflection,” and “knowing.” The latter is important but rarer indicators of adaptation and a sense of self that is necessary for resilience.

The researchers attributed topics to resilience indicators deductively, considering each topic’s top 15 words and associated top 20 posts. No assumptions were made regarding whether topics fit all resilience indicators and whether topics could be attributed to 1, >1, or no indicator. Strong agreement was achieved among the researchers, with some minor differences noted and discussed before an agreement was reached. The results of mapping topics to resilience indicators are shown in Table 2, with example posts.

Table 2. Result of mapping topics to resilience indicators.

Resilience indicator	Mapped topic	Example posts
Social capital	Gratitude, appreciation, daily routine, loving, supporting, well wishing, empathizing, fun, and friendship	<ul style="list-style-type: none"> “[...] you did good???? so you did do the dishes eventually. Just organised my bedside table and tidied backyard (courtyard), will get back to it in the morn, ta” [Gratitude] “Wooooohhhoooo thank you incredible human beings! Echoing x1000 [x’s] sentiment here” [Appreciation] “The way to get help out of those cycles is to talk about them. Would it help to write them in a piece of paper and hand that over to the pdoc?” [Supporting]
Belonging	Gratitude, appreciation, daily routine, loving, well wishing, celebrations, empathizing, welcoming, fun, and friendship	<ul style="list-style-type: none"> “Another welcome to the forum [...], it sounds like you’ve been through a lot so it’s understandable that you’re under so much stress.” [Welcoming] “Happy New Year everyone in NSW, Vic, ACT and Tassie” [Celebrations] “I get tired of my physical problems too but one step at a time my awesome friend, walking with you I am here for you all the way and so is [...]” [Friendship]
Learning	Reflection, supporting, and knowing	<ul style="list-style-type: none"> “Listening to music does help. Need to learn to dismiss those unhelpful negative voices.” [Supporting] “I’m trying so hard to see through my tears. it’s hard but I am determined to get there. Just knowing that you are here gives me strength x????” [Knowing]
Adaptive capacity	Reflection	<ul style="list-style-type: none"> “Will be thinking of you through this process [...] and pray that the transition is as smooth as possible.” [Reflection]
Self-efficacy	Facing difficulties and sleep	<ul style="list-style-type: none"> “Hi, I’m feeling very overwhelmed with anxiety atm. I’m trying to control it, but it feels like I’m a big shaking mess inside. I feel short of breath and my heart is racing” [Facing difficulties] “Anxiety is off the charts. I’m sleep deprived as I got zero sleep last night and my mood is extremely low.” [Sleep]

Most topics were attributed to the indicators “Social Capital” and “Belonging.” This aligns with the well-understood attributes of social media, which encourage and function through phatic communication, more so than informational or dialogic intents [53]. Within this broad sociopragmatic function of posts, we could see a range of topics that distinguished different methods of establishing and maintaining social connections. Topics associated with “Learning,” “Adaptive Capacity,” and “Self-efficacy” were rare, and these can be considered more explicit expressions of exhibited resilience, where change, adaptation, and coping become possible.

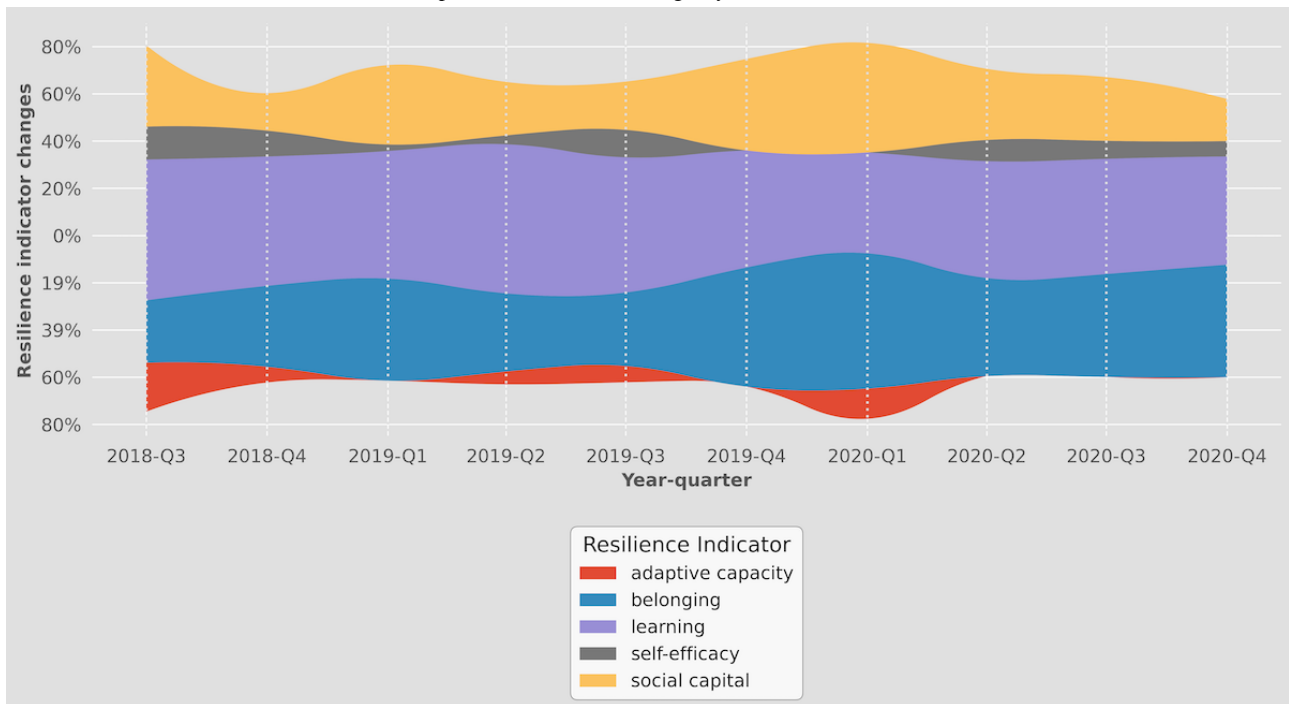
Results of Resilience Prevalence Analysis

Figure 7 shows the resilience streamgraph that resulted from the resilience prevalence analysis, which identified the dynamics of the realization of resilience indicators over time. This streamgraph was created to visualize our understanding of resilience prevalence [54], which is a popular method for displaying changes in different categories (ie, resilience indicators) over time. Instead of visualizing values as a conventional y-axis, streamgraphs offset the baseline of each “stack” to make it symmetrical around the x-axis. This results in a stream shape that illustrates the change in values over time. Through this streamgraph, we analyzed the patterns of resilience

indicators in terms of their realization over different quarters represented by the x-axis.

A total of 3 findings can be summarized in Figure 7. First, the prevalence of each indicator changes according to the proportion of posts (y-axis), with the height of each individual stream shape showing the prevalence of each indicator over time. The y-axis is not positive or negative but rather shows the best stacking arrangement. Second, “learning” and “belonging” were the top 2 most realized indicators over the entire timeline. The next most common indicator was “social capital,” whereas “self-efficacy” and “adaptive capacity” were less realized during the timeline, with shallow stream shapes appearing and disappearing at different points in time. Third, there were no distinct seasonal peaks and periodic patterns for the top 3 dominant indicators (ie, “learning,” “belonging,” and “social capital”) over the timeline. However, both “belonging” and “social capital” were observed to be slightly stronger at 2020-Q1. By contrast, the realization of both “self-efficacy” and “adaptive capacity” noticeably changed over time: (1) “self-efficacy” was hardly realized at the periods 2019-Q1, 2019-Q2, 2019-Q4, and 2020-Q1, and (2) “adaptive capacity” also seemed to be very weakly realized at the periods 2019-Q1, 2019-Q4, and 2020-Q2 to 2020-Q4. Investigating the possible reasons for these observations may be a useful area for future research.

Figure 7. Resilience indicator realization in Schizophrenia: A National Emergency (SANE) Australia.



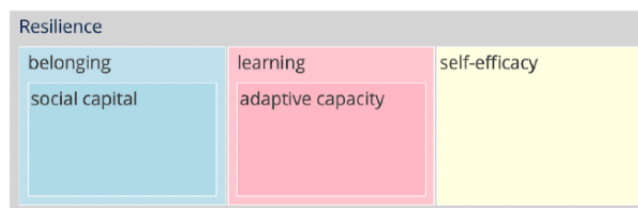
Phase 4: Resilience Dictionary Building

Result of the Resilience Taxonomy Inference

Given the co-occurrences of the resilience indicators mapped to the posts, we constructed a resilience taxonomy. We used a treemap representation (Figure 8) that shows the corealization of resilience indicators according to 2 levels. The first level has the indicators “belonging” and “learning.” These are higher-level and more general indicators. This means that these indicators were realized at a high prevalence, serving as foundational aspects of resilience. The indicator “social capital” is the more specific indicator subsumed by “belonging.” This means that whenever “social capital” is realized, “belonging” is most likely

to be realized as well. The same association was applied to “learning” and “adaptive capacity.” From another perspective, we can also draw an interpretation relating to the coverage of the 5 resilience indicators, where the coverage of each indicator is indicated by rectangular sizes. “Social capital” is clearly realized in relation to “belonging,” and “adaptive capacity” is realized in relation to “learning.” Colors are used to indicate similar characteristics, where more specific indicators are represented by darker intensities of the same colors. The indicator “self-efficacy” was realized independently, regardless of the other 4 indicators. The emergent taxonomy can be used to gain insights into which indicators are more likely to corealize and what types of conceptual relations are observable between indicators, as inferred by their corealization.

Figure 8. The resilience taxonomy induced from the SANE (Schizophrenia: A National Emergency) Australia posts represented as a treemap that shows the coverage of the concept of each resilience indicator by its rectangle size.



Result of Resilience Dictionary Generation

Table 3 shows the created resilience dictionary, demonstrating (1) the parent of each resilience indicator represented by “Parent,” (2) topics relevant to each indicator obtained from Table 1 by “Topic,” (3) the relative importance of topics by “Topic Weight,” and 4) the top 10 descriptive terms of each topic by “Topic Words.” The 5 most similar terms for each descriptive term are presented in Multimedia Appendix 1.

In summary, 2 sets of insights were provided by the resilience dictionary. First, sets of topics associated with different

resilience indicators were identified across the entire study period and within each quarter, revealing both constant and periodic topics. These topics and topic words form a useful set of semantic collections that can help characterize each resilience indicator. Second, topic weights, prevalence, and co-occurrences of topics associated with resilience indicators can be used to construct a resilience taxonomy. This taxonomy revealed foundational indicators of resilience—belonging and learning—and more specific and less prevalent indicators—social capital (subsumed by belonging) and adaptive capacity (subsumed by learning)—with self-efficacy realized independently of other indicators.

Table 3. The resilience dictionary constructed from this study.

Resilience indicator, parent, and topic	Topic weight (%)	Topic words
Social capital		
Belonging		
Gratitude	3.6	good sound morning news idea hear luck glad wish care
Appreciation	6.1	thank share appreciate reply word kind tag really look ur
Daily routine	7.7	day today photo walk birthday great nice week little yesterday
Loving	3.3	love hug send thinking cat care hon big wish hop
Supporting	2.4	help need support talk really care place try gp reach
Well wishing	1.6	hope ok birthday hey soon happy today enjoy nice glad
Empathizing	1.3	sorry hear really glad know sound care hard moment reach
Fun	1.6	lol ok work yeah today walk week yes maybe home
Friendship	1	friend hello care hug xx send awesome today big mum
Belonging		
Gratitude	3.6	good sound morning news idea hear luck glad wish care
Appreciation	6.1	thank share appreciate reply word kind tag really look ur
Daily routine	7.7	day today photo walk birthday great nice week little yesterday
Loving	3.3	love hug send thinking cat care hon big wish hop
Well wishing	1.6	hope ok birthday hey soon happy today enjoy nice glad
Celebrations	3.1	year happy new christmas birthday merry wish come last family
Empathizing	1.3	sorry hear really glad know sound care hard moment reach
Welcoming	11.5	support welcome member great community mental share health look experience
Fun	1.6	lol ok work yeah today walk week yes maybe home
Friendship	1	friend hello care hug xx send awesome today big mum
Learning		
Reflection	5.2	think people life way make try say need come maybe
Supporting	2.4	help need support talk really care place try gp reach
Knowing	44.8	know say want work people really hard let make life
Adaptive capacity		
Learning		
Reflection	5.2	think people life way make try say need come maybe
Self-efficacy		
Facing difficulties	2.6	feel really feeling make way free bad right moment pain
Sleep	4	sleep night tomorrow last tonight bed today work hon early

Discussion

Principal Findings

This paper presents a semiautomatic approach for generating a resilience dictionary from online mental health peer-support forum data. It applies a systematic 4-phase analysis pipeline (Figure 1) that preprocesses raw forum posts, discovers core themes, conceptualizes resilience indicators, and generates a resilience dictionary. We show the promise of exploring how resilience indicators are realized in a web-based mental health community to enrich the characterization of each resilience indicator by showing its range of topics or ways it is discussed

and the interdependencies between different aspects of resilience.

A major contribution of this study is that it provides a replicable method for generating further resilience and other topics or theory-focused dictionaries. The proposed method pipeline can be applied to online support forums hosted by other mental health organizations to better understand their web-based communities and build a deeper understanding of the realization of resilience.

To address RQ1, we present a method for using both the 2-layered NMF topic modeling technique and human knowledge through thematic analysis to discover and label the 15

substantive topics and map them to the 5 resilience indicators. Each resilience indicator was characterized by its relevant topics (Tables 2 and 3). The analysis showed that we discovered topics related to establishing and maintaining social connections and bonding, building trust, and belonging over time, including through the sharing of milestone events such as Christmas, New Year's Eve, and birthdays, as well as more routine moments. Reflection and knowledge sharing illustrate the importance of learning. In addition, sharing difficulties and ongoing issues associated with mental ill health such as lack of sleep illustrates the role of the forums for participants in navigating self-efficacy and agency amidst the mental health difficulties they face in their daily lives.

We traced the proportion of relevant posts over time to reveal the prevalence of the different aspects of resilience (RQ2). Figure 7 shows the dominance of "learning," "belonging," and "social capital," with "self-efficacy" and "adaptive capacity" appearing less frequently. These findings also help answer RQ3, addressing the relationship between different resilience indicators. To further clarify, we developed a resilience dictionary that shows how each resilience indicator is characterized through a set of relevant topics, the most relevant descriptive terms (including their synonyms) for each topic, and the semantic relationships between resilience indicators observed from the input post data (Table 3 and Multimedia Appendix 1). A taxonomy was developed to establish the relationships of interdependence between different resilience indicators (Figure 8), emphasizing the relationship between generators (belonging and learning) and outcome-oriented (social capital, adaptive capacity, and self-efficacy) indicators of resilience, in line with existing resilience research [27].

A growing body of research has sought to assess the impact of web-based mental health peer support. Peer-support forums offer a form of health intervention that can provide support at scale; however, they also present challenges regarding their design to optimize benefits and their evaluation. New computational methods for NLP, combined with qualitative analysis, provide opportunities to address these challenges. We leveraged NLP-based statistical analysis and qualitative analyses to generate new insights. This work extends beyond the capacity of qualitative content analysis studies that have attempted to identify the effects on resilience through surveys or interviews, cross-sectionally and at a small scale [9,14,18,19,21].

Our work identifies strengths-based indicators of resilience and the relationships between them. This represents an advance or change over previous studies that tended to focus on risk and negative issues regarding mental health, including identifying negative symptoms [31,32], emotions [33,34], and potential risks of self-harm [23]. Although responding to risk and intervening to support people is important, so is the capacity to look at the strengths and resources generated through web-based forums. Complementing these approaches, our work takes a strengths-based approach, with the resultant resilience dictionary able to add a positive context to deepen the understanding of the forum impact and its most effective attributes.

The benefits of the resilience dictionary are that it can be used for (1) showing that resilience, as understood by the current

theory, is realized through activity on forums; (2) improving understanding of the aspects of resilience realized on web-based forums; (3) building evidence-based resources for training staff and volunteers by showing how the different aspects of resilience are characterized and how resilience topics are embedded in certain discussions; (4) improving the design and automated moderation of mental health forums by enabling analysis to move beyond risk and harm to strengths-based indicators of resilience; and (5) aiding further text analysis as a vocabulary set for building resilience detection systems that can help mental health care providers to design to improve the outcomes of their services and illustrate their impact.

This study is significant for research and practice in potentially far-reaching areas. First, resilience itself, how it forms, and through what processes and means is poorly understood. Resilience tends to be an ill-defined concept in mental health practice. Here, we have shown how indicators, as part of a conceptual frame, resonate with activity and discussion over time in a web-based forum. Thus, by applying contemporary data analysis techniques to novel data, we have contributed to social and psychological knowledge about this contested concept. Second, for practice, this work shows the potential of (re)using novel data generated through a service to provide useful information that can help design and maintain those services, respond better to health consumers, and show their impact. These are all important issues for nonprofit organizations that struggle to secure and allocate scarce resources. This work should prompt mental health services to gear up their capacity to work with data as it shows considerable potential and innovative uses for data and data skills.

Limitations

This study has some limitations that could be addressed in future research. First, it focused on creating a resilience dictionary from a single forum and data source (ie, SANE Australia). It would be useful to build resilience dictionaries from multiple data sources and explore similarities and differences. In doing so, we might gain insights into common and distinctive resilience indicators across forums. If resilience realization is different, this could be explored in relation to varying forums and organizational aims, management strategies, and user demographics. Second, the study is "static" in the sense that it used a large existing corpus of historical data. Considering that web-based mental health forums are a growing type of service, it could be useful to investigate how to enable a resilience dictionary to incrementally evolve in response to new forum data.

Conclusions

In this paper, we presented a promising approach to creating a resilience dictionary that provides insights into the dominant topics on mental health peer-support forums and the way these topics realize different indicators of resilience. The developed mixed NLP and qualitative methods forge new grounds in helping forum providers analyze and understand the impact that the forums provide through a strengths-based analysis of resilience. Along with illustrating the prevalence of different resilience indicators over time, a taxonomy demonstrates the interdependence of different indicators, revealing which are

foundational (belonging and learning) in relation to others (social capital and adaptive capacity). The resulting resilience dictionary offers a benchmark and vocabulary set that can aid further research. It can also be used to inform automated systems

normally predicated on terms associated with risk, harm, and diagnostic indicators of mental ill health or distress, adding a strengths-based approach to the moderation of forum content and interactions.

Acknowledgments

This work was funded by the Australian Research Council Centre of Excellence for Automated Decision Making and Society (CE200100005, 2020-2027), and the Australian Research Council Discovery Project (DP200100419, 2020-2023). The authors would also like to thank Dr Sherridan Emery, University of Tasmania, for her assistance with the topic labeling.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Full resilience dictionary.

[[PDF File \(Adobe PDF File\), 91 KB-Multimedia Appendix 1](#)]

References

1. Mental health: strengthening our response. World Health Organization. 2022 Jun 17. URL: <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response> [accessed 2022-06-17]
2. Anderson K, Priebe S. Concepts of resilience in adolescent mental health research. *J Adolesc Health* 2021 Nov;69(5):689-695. [doi: [10.1016/j.jadohealth.2021.03.035](https://doi.org/10.1016/j.jadohealth.2021.03.035)] [Medline: [34045094](https://pubmed.ncbi.nlm.nih.gov/34045094/)]
3. Moßburger L, Wende F, Brinkmann K, Schmidt T. Exploring online depression forums via text mining: a comparison of Reddit and a curated online forum. In: Proceedings of the 5th Social Media Mining for Health Applications Workshop & Shared Task. 2020 Presented at: SMM4H '20; June 28, 2020; Barcelona, Spain p. 70-81 URL: <https://aclanthology.org/2020.smm4h-1.11> [doi: [10.18653/v1/w18-5906](https://doi.org/10.18653/v1/w18-5906)]
4. Park I, Sarnikar S, Cho J. Disentangling the effects of efficacy-facilitating informational support on health resilience in online health communities based on phrase-level text analysis. *Inf Manag* 2020 Dec;57(8):103372 [FREE Full text] [doi: [10.1016/j.im.2020.103372](https://doi.org/10.1016/j.im.2020.103372)]
5. Ritchie H. Global mental health: five key insights which emerge from the data. Our World In Data. 2018 May 16. URL: <https://ourworldindata.org/global-mental-health> [accessed 2021-09-16]
6. Cook L. Mental health in Australia: a quick guide. Parliament of Australia. 2019 Feb 14. URL: https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1819/Quick_Guides/MentalHealth [accessed 2021-09-16]
7. National Study of Mental Health and Wellbeing. First insights from the National Study of Mental Health and Wellbeing, 2020-21. Australian Bureau of Statistics. 2021 Aug 12. URL: <https://www.abs.gov.au/articles/first-insights-national-study-mental-health-and-wellbeing-2020-21> [accessed 2022-04-13]
8. Chan AC, Piehler TF, Ho GW. Resilience and mental health during the COVID-19 pandemic: findings from Minnesota and Hong Kong. *J Affect Disord* 2021 Dec 01;295:771-780 [FREE Full text] [doi: [10.1016/j.jad.2021.08.144](https://doi.org/10.1016/j.jad.2021.08.144)] [Medline: [34517251](https://pubmed.ncbi.nlm.nih.gov/34517251/)]
9. Verdolini N, Amoretti S, Montejó L, García-Rizo C, Hogg B, Mezquida G, et al. Resilience and mental health during the COVID-19 pandemic. *J Affect Disord* 2021 Mar 15;283:156-164 [FREE Full text] [doi: [10.1016/j.jad.2021.01.055](https://doi.org/10.1016/j.jad.2021.01.055)] [Medline: [33556749](https://pubmed.ncbi.nlm.nih.gov/33556749/)]
10. Mayo Clinic Staff. Resilience: Build skills to endure hardship. Mayo Clinic. URL: <https://www.mayoclinic.org/tests-procedures/resilience-training/in-depth/resilience/art-20046311> [accessed 2021-09-17]
11. Depression - Beyond Blue Forums. URL: <https://www.beyondblue.org.au/get-support/online-forums/depression> [accessed 2022-06-01]
12. SANE Australia. URL: <https://www.sane.org/> [accessed 2022-06-01]
13. ReachOut. URL: <https://au.reachout.com/> [accessed 2022-09-07]
14. Hechanova MR, Alianan Jr AS, Jocson RM, Villasanta AP. Online resilience support groups during the COVID-19 pandemic: the Philippine experience. *Ment Health Soc Inclusion* 2022;26(1):56-65. [doi: [10.1108/mhsi-06-2021-0038](https://doi.org/10.1108/mhsi-06-2021-0038)]
15. Fisher EB, Miller SM, Evans M, Luu SL, Tang PY, Dreyer Valovcin D, et al. COVID-19, stress, trauma, and peer support-observations from the field. *Transl Behav Med* 2020 Aug 07;10(3):503-505 [FREE Full text] [doi: [10.1093/tbm/ibaa056](https://doi.org/10.1093/tbm/ibaa056)] [Medline: [32569372](https://pubmed.ncbi.nlm.nih.gov/32569372/)]
16. McCosker A. Engaging mental health online: insights from beyondblue's forum influencers. *New Media Soc* 2018 Jul 11;20(12):4748-4764 [FREE Full text] [doi: [10.1177/1461444818784303](https://doi.org/10.1177/1461444818784303)]

17. Gillard S, Foster R, Gibson S, Goldsmith L, Marks J, White S. Describing a principles-based approach to developing and evaluating peer worker roles as peer support moves into mainstream mental health services. *Ment Health Soc Inclusion* 2017 Jun 12;21(3):133-143. [doi: [10.1108/mhsl-03-2017-0016](https://doi.org/10.1108/mhsl-03-2017-0016)]
18. Wu Y, Sang ZQ, Zhang XC, Margraf J. The relationship between resilience and mental health in Chinese college students: a longitudinal cross-lagged analysis. *Front Psychol* 2020 Feb 5;11:108 [FREE Full text] [doi: [10.3389/fpsyg.2020.00108](https://doi.org/10.3389/fpsyg.2020.00108)] [Medline: [32116918](https://pubmed.ncbi.nlm.nih.gov/32116918/)]
19. Riehm KE, Brenneke SG, Adams LB, Gilan D, Lieb K, Kunzler AM, et al. Association between psychological resilience and changes in mental distress during the COVID-19 pandemic. *J Affect Disord* 2021 Mar 01;282:381-385 [FREE Full text] [doi: [10.1016/j.jad.2020.12.071](https://doi.org/10.1016/j.jad.2020.12.071)] [Medline: [33421866](https://pubmed.ncbi.nlm.nih.gov/33421866/)]
20. Building your resilience. American Psychological Association. URL: <https://www.apa.org/topics/resilience> [accessed 2021-09-16]
21. Liu WJ, Zhou L, Wang XQ, Yang BX, Wang Y, Jiang JF. Mediating role of resilience in relationship between negative life events and depression among Chinese adolescents. *Arch Psychiatr Nurs* 2019 Dec;33(6):116-122. [doi: [10.1016/j.apnu.2019.10.004](https://doi.org/10.1016/j.apnu.2019.10.004)] [Medline: [31753216](https://pubmed.ncbi.nlm.nih.gov/31753216/)]
22. Conway M, Hu M, Chapman WW. Recent advances in using natural language processing to address public health research questions using social media and ConsumerGenerated data. *Yearb Med Inform* 2019 Aug;28(1):208-217 [FREE Full text] [doi: [10.1055/s-0039-1677918](https://doi.org/10.1055/s-0039-1677918)] [Medline: [31419834](https://pubmed.ncbi.nlm.nih.gov/31419834/)]
23. Milne DN, McCabe KL, Calvo RA. Improving moderator responsiveness in online peer support through automated triage. *J Med Internet Res* 2019 Apr 26;21(4):e11410 [FREE Full text] [doi: [10.2196/11410](https://doi.org/10.2196/11410)] [Medline: [31025945](https://pubmed.ncbi.nlm.nih.gov/31025945/)]
24. r/depression stats. Reddit. URL: <https://subredditstats.com/r/depression> [accessed 2022-06-01]
25. HealthBoards. URL: <https://www.healthboards.com> [accessed 2022-06-01]
26. Inspire. URL: <https://www.inspire.com> [accessed 2022-06-01]
27. HealthUnlocked. URL: <https://www.healthunlocked.com> [accessed 2022-06-01]
28. Davcheva E. Text mining mental health forums - learning from user experiences. In: Proceedings of the 26th European Conference on Information Systems. 2018 Presented at: ECIS '18; June 23-28, 2018; Portsmouth, UK p. 91 URL: https://aisel.aisnet.org/ecis2018_rp/91/
29. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, et al. Global sentiments surrounding the COVID-19 pandemic on Twitter: analysis of Twitter trends. *JMIR Public Health Surveill* 2020 May 22;6(2):e19447. [doi: [10.2196/19447](https://doi.org/10.2196/19447)] [Medline: [32412418](https://pubmed.ncbi.nlm.nih.gov/32412418/)]
30. Calvo RA, Milne DN, Hussain MS, Christensen H. Natural language processing in mental health applications using non-clinical texts. *Nat Lang Eng* 2017 Jan 30;23(5):649-685. [doi: [10.1017/S1351324916000383](https://doi.org/10.1017/S1351324916000383)]
31. Park A, Conway M. Longitudinal changes in psychological states in online health community members: understanding the long-term effects of participating in an online depression community. *J Med Internet Res* 2017 Mar 20;19(3):e71 [FREE Full text] [doi: [10.2196/jmir.6826](https://doi.org/10.2196/jmir.6826)] [Medline: [28320692](https://pubmed.ncbi.nlm.nih.gov/28320692/)]
32. Yazdavar AH, Mahdavinejad MS, Bajaj G, Romine W, Sheth A, Monadjemi AH, et al. Multimodal mental health analysis in social media. *PLoS One* 2020 Apr 10;15(4):e0226248 [FREE Full text] [doi: [10.1371/journal.pone.0226248](https://doi.org/10.1371/journal.pone.0226248)] [Medline: [32275658](https://pubmed.ncbi.nlm.nih.gov/32275658/)]
33. Patton DU, MacBeth J, Schoenebeck S, Shear K, McKeown K. Accommodating grief on Twitter: an analysis of expressions of grief among gang involved youth on Twitter using qualitative analysis and natural language processing. *Biomed Inform Insights* 2018 Apr 3;10:1178222618763155 [FREE Full text] [doi: [10.1177/1178222618763155](https://doi.org/10.1177/1178222618763155)] [Medline: [29636619](https://pubmed.ncbi.nlm.nih.gov/29636619/)]
34. Loveys K, Crutchley P, Wyatt E, Coppersmith G. Small but mighty: affective micropatterns for quantifying mental health from social media language. In: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology. 2017 Presented at: CLPsych '17; August 3, 2017; Vancouver, Canada p. 85-95. [doi: [10.18653/v1/w17-3110](https://doi.org/10.18653/v1/w17-3110)]
35. SANE Australia Annual Highlights 2019-2020. SANE Australia. 2021. URL: <https://www.sane.org/about-sane/annual-reports-highlights> [accessed 2022-06-01]
36. Berkes F, Ross H. Community resilience: toward an integrated approach. *Soc Natural Resources* 2013 Jan;26(1):5-20 [FREE Full text] [doi: [10.1080/08941920.2012.736605](https://doi.org/10.1080/08941920.2012.736605)]
37. Brown K, Westaway E. Agency, capacity, and resilience to environmental change: lessons from human development, well-being, and disasters. *Annu Rev Environ Resour* 2011 Nov 21;36(1):321-342 [FREE Full text] [doi: [10.1146/annurev-environ-052610-092905](https://doi.org/10.1146/annurev-environ-052610-092905)]
38. Skerratt S. Enhancing the analysis of rural community resilience: evidence from community land ownership. *J Rural Stud* 2013 Jul;31:36-46 [FREE Full text] [doi: [10.1016/j.jrurstud.2013.02.003](https://doi.org/10.1016/j.jrurstud.2013.02.003)]
39. Fazey I, Fazey JA, Fischer J, Sherren K, Warren J, Noss RF, et al. Adaptive capacity and learning to learn as leverage for social-ecological resilience. *Front Ecol Environ* 2007 Sep;5(7):375-380 [FREE Full text] [doi: [10.1890/1540-9295\(2007\)5\[375:acaltl\]2.0.co;2](https://doi.org/10.1890/1540-9295(2007)5[375:acaltl]2.0.co;2)]
40. Reisenbichler M, Reutterer T. Topic modeling in marketing: recent advances and research opportunities. *J Bus Econ* 2019;89(3):327-356. [doi: [10.1007/s11573-018-0915-7](https://doi.org/10.1007/s11573-018-0915-7)]
41. Albalawi R, Yeap TH, Benyoucef M. Using topic modeling methods for short-text data: a comparative analysis. *Front Artif Intell* 2020 Jul 14;3:42 [FREE Full text] [doi: [10.3389/frai.2020.00042](https://doi.org/10.3389/frai.2020.00042)] [Medline: [33733159](https://pubmed.ncbi.nlm.nih.gov/33733159/)]

42. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature* 1999 Oct 21;401(6755):788-791. [doi: [10.1038/44565](https://doi.org/10.1038/44565)] [Medline: [10548103](https://pubmed.ncbi.nlm.nih.gov/10548103/)]
43. O'Callaghan D, Greene D, Carthy J, Cunningham P. An analysis of the coherence of descriptors in topic modeling. *Expert Syst Appl* 2015 Aug;42(13):5645-5657 [FREE Full text] [doi: [10.1016/j.eswa.2015.02.055](https://doi.org/10.1016/j.eswa.2015.02.055)]
44. Aizawa A. An information-theoretic perspective of tf-idf measures. *Inf Process Manag* 2003 Jan;39(1):45-65 [FREE Full text] [doi: [10.1016/s0306-4573\(02\)00021-3](https://doi.org/10.1016/s0306-4573(02)00021-3)]
45. Belford M, Mac Namee B, Greene D. Stability of topic modeling via matrix factorization. *Expert Syst Appl* 2018 Jan;91(C):159-169. [doi: [10.1016/j.eswa.2017.08.047](https://doi.org/10.1016/j.eswa.2017.08.047)]
46. Isoaho K, Gritsenko D, Mäkelä E. Topic modeling and text analysis for qualitative policy research. *Policy Stud J* 2021 Feb;49(1):300-324 [FREE Full text] [doi: [10.1111/psj.12343](https://doi.org/10.1111/psj.12343)]
47. Kang YB, Haghhigh PD, Burstein F. TaxoFinder: a graph-based approach for taxonomy learning. *IEEE Trans Knowl Data Eng* 2016 Feb 1;28(2):524-536. [doi: [10.1109/tkde.2015.2475759](https://doi.org/10.1109/tkde.2015.2475759)]
48. Kang YB, Woo J, Kneebone L, Sellis T. Methodology for refining subject terms and supporting subject indexing with taxonomy: a case study of the APO digital repository. *Decis Support Syst* 2021 Jul;146:113542 [FREE Full text] [doi: [10.1016/j.dss.2021.113542](https://doi.org/10.1016/j.dss.2021.113542)]
49. de Knijff J, Frasinca F, Hogenboom F. Domain taxonomy learning from text: the subsumption method versus hierarchical clustering. *Data Knowl Eng* 2013 Jan;83:54-69 [FREE Full text] [doi: [10.1016/j.datak.2012.10.002](https://doi.org/10.1016/j.datak.2012.10.002)]
50. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. *Comput Res Repository* 2013 [FREE Full text]
51. Chen Y, Zhang H, Liu R, Ye Z, Lin J. Experimental explorations on short text topic mining between LDA and NMF based schemes. *Knowl Based Syst* 2019 Jan;163:1-13 [FREE Full text] [doi: [10.1016/j.knosys.2018.08.011](https://doi.org/10.1016/j.knosys.2018.08.011)]
52. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
53. Miller V. New media, networking and phatic culture. *Convergence* 2008 Nov 01;14(4):387-400. [doi: [10.1177/1354856508094659](https://doi.org/10.1177/1354856508094659)]
54. Bider L. Streamgraphs: how to make them and what you need to know. *Flourish*. 2020 Oct 19. URL: <https://flourish.studio/2020/10/19/streamgraphs/> [accessed 2021-10-10]

Abbreviations

- NLP:** natural language processing
NMF: nonnegative matrix factorization
RQ: research question
SANE: Schizophrenia: A National Emergency

Edited by T Leung; submitted 25.04.22; peer-reviewed by S Chen, G Makwana; comments to author 19.05.22; revised version received 06.06.22; accepted 18.08.22; published 22.09.22

Please cite as:

Kang YB, McCosker A, Kamstra P, Farmer J

Resilience in Web-Based Mental Health Communities: Building a Resilience Dictionary With Semiautomatic Text Analysis

JMIR Form Res 2022;6(9):e39013

URL: <https://formative.jmir.org/2022/9/e39013>

doi: [10.2196/39013](https://doi.org/10.2196/39013)

PMID:

©Yong-Bin Kang, Anthony McCosker, Peter Kamstra, Jane Farmer. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 22.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.