







Building a Tool that Draws from the Collective Wisdom of the Internet to Help Users Respond Effectively to Anxiety-Related Questions

Benjamin T. Kaveladze¹ , George I. Kaveladze² , Elad Yom-Tov^{3,4} ,
and Stephen M. Schueller^{1,5} 

¹ Department of Psychological Science, University of California, Irvine, Irvine, CA, USA
bkavelad@uci.edu

² The MITRE Corporation, Bedford, MA, USA

³ Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology,
Haifa, Israel

⁴ Microsoft Research, Herzliya, Israel

⁵ Department of Informatics, University of California, Irvine, Irvine, CA, USA

Abstract. Online anxiety support communities offer a valuable and accessible source of informational and emotional support for people around the world. However, effectively responding to posters' anxiety-related questions can be challenging for many users. We present our work in developing a web-based tool that draws from previous question-response interactions and trusted online informational resources to help users rapidly produce high-quality responses to anxiety-related questions. We describe our efforts in four parts: 1) Creating a machine learning classifier to predict response quality, 2) developing and evaluating a computational question-answering system that learns from previous questions and responses on support forums, 3) developing and evaluating a system to suggest online resources for anxiety-related questions, and 4) interviewing support community moderators to inform further system design. We discuss how this tool might be integrated into online anxiety support communities and consider challenges with the tool's functionality and implementation. We also provide the dataset we used to train the system to provide opportunities for other researchers to build on this work.

Keywords: Online mental health communities · Question answering · Big data · Online support provision · Anxiety

1 Introduction

1.1 Questions and Answers on Online Mental Health Communities

Mental health struggles are extremely common globally, but many people lack reliable access to professional mental healthcare or sensitive social networks from which they can draw support (Kohn et al. 2004). Online mental health support communities offer accessible opportunities for people to exchange advice and validation with others

who understand or share their struggles. These spaces take advantage of availability, anonymity, and ease of use to bypass common barriers to receiving mental health support, such as stigma, shyness, and physical isolation (Bargh 2002). As such, they can be especially crucial resources for young people or members of marginalized groups who might be concerned about the social consequences of seeking help from their in-person networks (O’Leary et al. 2017, Rains and Tsetsi 2017).

Subreddits are free, user-led online discussion forums hosted by Reddit.com. There are many subreddits related to mental health, each with its own subculture and unique forms of support (Sharma and De Choudhury 2018), but in the present work we focus on anxiety-related subreddits. While the norm of using Reddit anonymously precludes obtaining representative demographic information on subreddit users, most respondents to a survey of mental health subreddit users conducted in December 2019 ($n > 300$) were young (50% under 24), white (79%), female (56%), and American (59%, though respondents came from 44 countries in total); further, most rated their mental health as terrible or poor (82%) and had received professional mental healthcare in the past (87%) (Kaveladze and Schueller 2020).

Mental health subreddit users share a range of content, including uplifting messages, emotional “vents”, and tips. However, many interactions across online mental health communities involve a user posting to ask for help with a personal problem (Kaveladze and Schueller 2020). After a question is submitted, other members comment on that submission to offer advice or support, typically within a few minutes to a few days later (De Choudhury and De 2014). Knowing how to respond effectively to a stranger’s question about a private mental health issue can be challenging. As a result, many questions and support requests on these spaces go unanswered or receive unhelpful answers, and negativity and misinformation in responses are common issues (Kaveladze and Schueller 2020).

While every question is unique, we observe that many anxiety-related questions are similar to other questions that received effective responses in the past. We further observe that many of these responses are accessible via the internet, either in publicly-archived question-response interactions from online support communities or in online informational resources. Indeed, over 15,000 such question-response interactions occur annually on one popular anxiety subreddit. Thus, the internet provides a corpus of collective wisdom on responding to anxiety-related struggles that is unprecedented in its size and the diversity of its contributors. We aim to build an automated tool that leverages this corpus to help users respond to new anxiety-related questions. Specifically, the tool aims to help users craft high-quality and well-informed supportive responses quicker than they would otherwise.

1.2 Accessing Collective Wisdom from Mental Health Subreddits

Subreddits are particularly amenable to computational analyses because all posts are publicly available and because the Reddit API and interconnected data download tools like Pushshift enable large-scale downloads of post data (Baumgartner et al. 2020). Several studies have used computational approaches to study support interactions in online mental health communities. The majority of this research has examined help-seeking posts, identifying factors such as linguistic accommodation that predict posts’ tendencies

to receive responses expressing emotional and informational support (De Choudhury and De 2014; Sharma and De Choudhury 2018; Gkotsis et al. 2016). However, some research has explored the characteristics of helpful responses to posts on online mental health communities. Peng et al. (2021) found that mental health subreddit posters responded more positively to responses to their help-seeking posts that matched the kind of support they requested (informational vs. emotional). Also, Kavuluru et al. (2016) developed a computational model to predict responses' helpfulness from a suicide prevention perspective. Their model showed promising results, but it only included a few predictors from the response text and left out post metadata such as time to respond. We build on Kavuluru et al.'s approach, including a wider range of linguistic and metadata predictors aiming to create a more comprehensive model of response quality.

1.3 Project Summary

We describe the process of developing the helper tool in four parts, each with an overview, methods, and results section. Part 1 describes a machine learning classifier we built to label responses to anxiety-related questions as high- or low-quality. Part 2 describes a question-answering system we designed to take in anxiety-related questions and suggest responses pulled from a corpus of question-response interactions. Part 3 describes a system that matches questions to relevant online informational resources. Part 4 describes our current progress in designing the helper tool, including interviewing stakeholders and identifying challenges to the tool's effectiveness and implementation. This work's contribution is to inform efforts to improve supports for online mental health community members. In addition, we provide the dataset used to train our system for other researchers to learn from and improve on our method.

2 Creating the Helper Tool

2.1 Part 1: Creating a Machine Learning Classifier to Predict Response Quality

Overview. We drew from a corpus of 12,325 responses to 7,646 questions and associated metadata from two online support forums. Next, we gathered a subset of those responses and manually labeled the quality and type of support they offered. We then built a random forest classifier model (cross-validated AUC = 0.82) that labeled responses to anxiety-related questions as high- or low-quality.

Methods

Dataset. We examined two Anxiety-related subreddits that had been popular for several years (at the time of data collection, we observed that one subreddit often had over 1,000 subscribers logged into Reddit at any given time and the other often had over 250). Both of these subreddits required posters to tag their submissions with one of several flairs (subreddit-specific tags) describing their submission. We gathered submissions with the "Help" flair from one subreddit and with the "Advice needed" flair from the other subreddit as these flairs identified submissions asking questions. We also gathered comments associated with each of those posts. Next, we limited our dataset to comments

responding to the submission, rather than those responding to other comments, and to posts, both submissions and comments, with fewer than 250 words and more than 5 words. We chose this word range as posts with 5 words or fewer may have been too vague to interpret and posts with more than 250 words may have taken too long for human raters to read. With these inclusion criteria, we used Python to query the Pushshift API and the Reddit API, yielding a dataset of 12,325 question-response pairs and their associated metadata posted between July of 2017 and February of 2021 on these two subreddits (Rossum and Drake 2009).

In addition to question and response text, we obtained post-level metadata on the date and time of posting and post score. A post’s score is the sum of upvotes [+1] and downvotes [-1] that it receives – users with Reddit accounts can add one upvote or downvote to each post. The dataset also contained user-level metadata on karma (the sum of upvotes minus downvotes received across all of a user’s Reddit posts) and account formation date/time. Finally, we created several textual variables for each post, including sentiment, readability, and the proportion of words matching several Linguistic Inquiry and Word Count (LIWC) dictionaries (negative, positive, feeling, anxiety, health, affiliative drive, body, anger, and sadness-related words) (Tausczik and Pennebaker 2010). We calculated average sentiment at the post level using the AFINN sentiment measure (Nielsen 2011), which provides values from -5 (negative sentiment) to 5 (positive sentiment) for emotionally-valenced words.

Data Privacy. All post data we analyzed in this work were publicly accessible and thus exempt from IRB review at the University of California, Irvine. However, most subreddit users likely do not expect that their posts will be used for research, making data privacy especially important. To help protect user privacy, we do not mention any potentially identifiable information and we removed the post authors’ usernames and post text from the dataset we share.

Rating Response Quality. We selected a random sample of 790 question-response interactions from our dataset and recruited 365 crowdsourced workers from Amazon Mechanical Turk (Turkers) who self-identified as fluent English speakers to rate responses to 10 randomly chosen questions from the sample. Turkers were presented the question and response and asked to rate how well each response answered the question (not at all well [0] – very well [3]), whether the response provided emotional support (yes/no), and whether the response provided instrumental support (yes/no). Raters were told to imagine that they asked the question and received the response anonymously (Mazuz and Yom-Tov 2020). An example of a high-quality response to a mental health question is shown in Fig. 1.

To monitor the quality of the Turker ratings, we dropped rating data from Turkers who did at least one of the following: failed the attention check during the task, self-reported at the end that they just skipped through the task without reading, or gave the same rating to all 10 of the interactions they rated. Based on these exclusion criteria, we dropped ratings from 270/365 Turkers, leaving plausibly legitimate ratings from only 95 Turkers. To supplement the Turker ratings, we trained three research assistants to rate all 790

question-response interactions in the same way that the Turkers did. The research assistants provided reasonably consistent ratings (Krippendorff’s alphas: subjective response quality = 0.68, emotional support = 0.69, informational support = 0.75). However, research assistants’ and Turkers’ ratings were less consistent (Krippendorff’s alphas: subjective response quality = 0.54, emotional support = 0.36, informational support = 0.43). We transformed these ratings into a binary response quality variable, defining high-quality responses as those with ratings (averaged across all raters) above all three scale midpoints (1.5 for the subjective response quality scale, and 0.5 for the emotional support and informational support scales). Using this binary response quality variable, 399 responses were labeled as high-quality and 391 as low-quality.

Classification Model. We trained a random forest classification model (scikit-learn library, Pedregosa et al. 2011) on the human-rated dataset to predict if answers were of high quality. We used a random forest model because they are known to be the best overall models for structured data, together with SVMs and neural networks (see Fernández-Delgado et al. 2014 and the criticism thereof: Wainberg et al. 2016). The independent attributes of the model were syllable count, character count, word count, sentence count, readability (Flesch 1948), sentiment valence (Nielsen 2011), post score, percentage of words matching the LIWC “negative emotion”, “positive emotion”, “feeling”, “anxiety”, “health”, “affiliative drive”, “body”, “anger”, and “sadness” dictionaries (Nielsen 2011), posting date, and time of day of posting, all for both questions and responses. We also included question-asker and respondent account formation date and account karma as attributes. We used 10-fold cross-validation to assess the accuracy of the model.

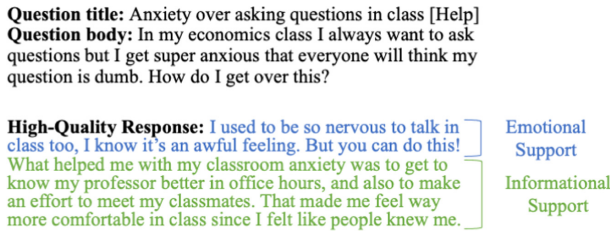


Fig. 1. An example of a high-quality response to an anxiety-related question, including emotional and informational support.

Results. In analyses, we used ordinary least squares regressions to detect linear relationships, Welch’s t-tests to compare group means while accounting for unequal variance across groups, and Spearman’s rank-order correlations (ρ) to track associations between non-normally distributed variables. All analyses were performed in the R statistical computing language using the “stats” package (R Core Team 2013). Data manipulation and figure creation used the “Tidyverse” family of packages in R (Wickham 2019).

Classification Model. Our cross-validated response quality classification model achieved a mean AUC of 0.82 and mean accuracy of 0.72. Each model feature’s importance in the model is shown in Fig. 2. We applied this model to classify the 11,535 responses that had not been rated by humans.

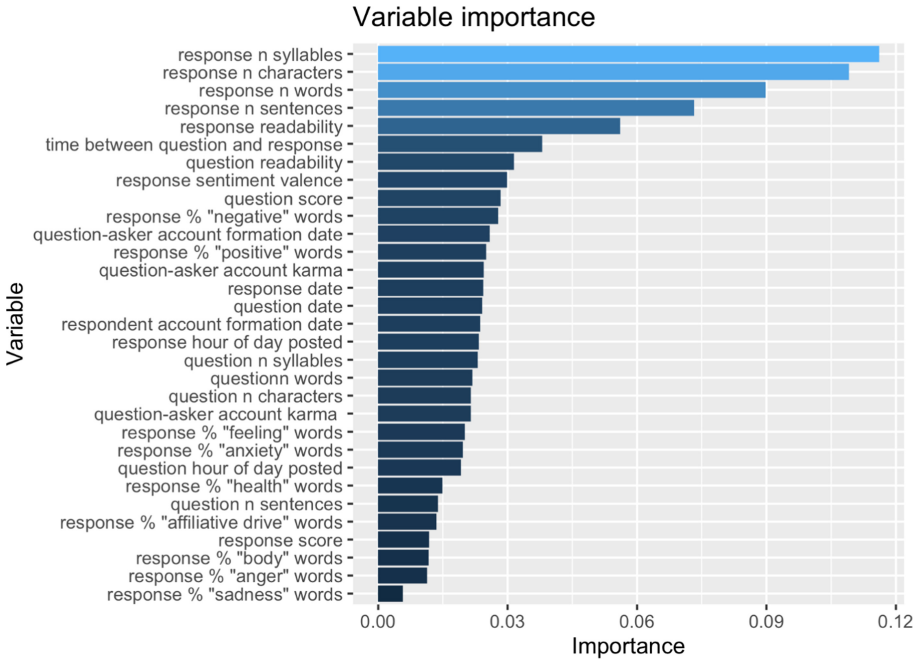


Fig. 2. Each feature’s importance in the random forest model predicting response quality. Lighter blue signifies higher importance. Importance (GINI importance) was computed as the (normalized) total reduction of the criterion brought by that feature. “n syllables” refers to the number of syllables in the question or response text. “readability” refers to the Flesch reading ease score calculated on the response text (Flesch 1948). Time between question and response is the length of time between when the question was posted and the response was posted. Sentiment valence refers to the sentiment score using the AFINN sentiment measure (Nielsen 2011). Score refers to the sum of “upvotes” (+1) and “downvotes” (−1) on a post. % “negative words” refers to the percentage of words in a post that matched words from the “negative emotion” Linguistic Inquiry and Word Count (LIWC) dictionary (Tausczik and Pennebaker 2010). Account formation date refers to the date and time that a user’s Reddit account was made. Account karma refers to the sum, across all a users’ posts in their account history, of their posts’ scores.

2.2 Part 2: Developing and Testing a Question-Answering System

Overview. We built on the predictive model from Part 1 to develop a question-answering (QA) system. Drawing from previous question-response interactions, the system aimed to match new anxiety-related questions to high-quality responses from previous interactions based on linguistic similarity. We then evaluated this tool’s effectiveness with help from crowdworkers and found that it outperformed chance but did not provide acceptable responses for a majority of new questions.

Methods. Building on the original dataset of 12,365 question-response pairs, we gathered another 4,272 question-response pairs from the two subreddits posted between

May of 2020 and May of 2021, including their associated metadata. Using this dataset of 16,637 pairs, we developed a question-answering system to identify previous high-quality responses that might be relevant to a novel question.

Identifying Useful Responses. To identify potentially useful responses, we filtered the question-response pairs using the response quality classifier we developed in part 1, keeping only responses that the classifier labeled as high-quality.

Finding Similar Responses. Working under the assumption that similar questions could receive similar answers, we attempted to identify useful answers to new questions by finding the most similar past question that had a response classified as high-quality. However, questions are often lengthy and contain information that is sometimes irrelevant to the answer. Therefore, instead of finding the textual similarity between questions, we first applied abstractive summarization to all questions and then found the similarity among summarized questions.

Specifically, questions were summarized to a predefined length of 50 words using the Hugging Face summarization pipeline [huggingface.co]. In our work we used a maximal summary length of 50 words. We evaluated other lengths between 35 and 75 words but did not observe significant differences in performance. We estimated the similarity between questions using cosine similarity to the TF-IDF normalized word and bigram representation of the summarized questions.

Evaluation. Next, we recruited 106 crowdsourced workers from Prolific to test the QA system’s effectiveness. Each participant viewed a series of responses to questions from the anxiety subreddits. Each of these responses was either 1) the original response to the question posted on Reddit, 2) a response to a question in our dataset that the QA system deemed was the closest, or 3) a randomly selected response to another question from our dataset. For each interaction, workers answered the following questions: “Does the response address the issues that the questions asked about? (y/n)” and “If you asked the question above, would you have been satisfied with the answer? (y/n)”. Due to duplicate responses, there were 69 original question-response interactions but only 67 randomly selected responses and 60 QA system-selected “closest high-quality” responses. We conducted one-way ANOVAs to compare the means between types of responses, averaged across all raters ($y = 1, n = 0$), and ran planned Tukey’s HSD tests to compare group means.

Results

Response Relevance and Satisfaction. The crowdsourced workers recruited from Prolific provided consistent ratings of response quality (Krippendorff’s alphas = 0.73 for both response relevance and satisfaction). Average response relevance ratings were lowest for randomly selected responses ($M = 0.18, SD = 0.38, n = 67$), higher for closest high-quality responses ($M = 0.26, SD = 0.44, n = 60$), and highest for original responses ($M = 0.52, SD = 0.50, n = 69$), $F = 55.88, p < 0.001$, and a Tukey’s HSD test found that all group means were significantly different (p ’s < 0.001). Response satisfaction ratings were lowest for randomly selected responses ($M = 0.24, SD = 0.43, n = 67$),

higher for closest high-quality responses ($M = 0.38$, $SD = 0.49$, $n = 60$), and highest for original responses ($M = 0.74$, $SD = 0.44$, $n = 69$) ($F = 116.20$, $p < 0.001$), and a Tukey's HSD test found that all group means were significantly different (p 's < 0.05).

2.3 Part 3: Developing a System to Suggest Online Resources for Anxiety-Related Question

Overview. To supplement the responses provided by the QA system, we developed a system to suggest relevant online resources for anxiety-related questions. The system recommended an appropriate resource 51% of the time, compared to 37% for a random choice.

Methods. We assembled a corpus of links to 80 online resources, 72 of which were from mhanational.org, a website providing free and accessible expert information about various mental health issues. The remaining 8 came from other informational websites. These resources provided information or other support for anxiety-related issues, such as advice on how to overcome social anxiety at parties or how to calm one's mind (corpus available online). A graduate student and undergraduate research assistant matched 250 questions from the dataset to an online resource from the corpus that they felt might be most helpful to the question-asker.

We developed a multiclass learning model to predict the appropriate online resources for particular questions. Our model ignored resources with fewer than 5 matched questions, leaving 26 online resources. The model was a Random Forest classifier with 10 trees and it used the TF-IDF representation of the words and bigrams of the questions to predict which of the information pages should be added to the answer. The model's performance was evaluated using 10-fold cross-validation.

Results. The online resource prediction model achieved 51% mean accuracy, compared to 37% mean accuracy for a model that did not use multiclass learning, and instead simply selected a resource for each question based only on the frequency with which that resource was chosen in the 250 training questions.

2.4 Part 4: Designing and Implementing the Helper Tool

Overview. We conducted interviews with moderators from mental health subreddits to gain stakeholder perspectives on how a tool using this QA system might be best designed and implemented. Some moderators expressed enthusiasm for the tool's potential to improve the quality and rapidity of responses, while other moderators were concerned that it could lead to less meaningful interactions in the community.

Methods. We reached out to nine moderators of mental health subreddits to invite them to participate in a semi-structured 30-min interview over video chat with one of the researchers on our team. During the interview, we walked moderators through a mockup of the tool (shown in Fig. 3) and asked them several questions regarding how helpful

they felt it could be for moderators and users, how it could be improved to be more useful for subreddit users, and how comfortable they would be with the tool being used on mental health subreddits.

Results

Qualitative Interviews. Two mental health subreddits moderators agreed to participate in a video interview. These moderators were comfortable with the idea of the tool as we presented it and were not concerned about data privacy issues. The moderators felt that such a tool would help both moderators and non-moderators – many of whom struggle with how to formulate effective responses to mental health questions – to more rapidly and easily provide high-quality responses to mental health questions. One moderator noted that if the tool was available to all users, it would enable users to screen their own questions before posting, helping to deal with the common problem of similar questions being posted repeatedly on a subreddit by different users. One moderator also noted that the Reddit search tool did not function well at finding past responses that were relevant to a new question, so the tool could fill that role. Both moderators suggested it would be useful to make the tool integrate directly with the Reddit API so that it could be easily used while browsing the site. Both moderators also stressed that our research team should stay in communication with the mental health communities to get their input on the tool throughout the development process.

Two other mental health subreddit moderators responded to our invitations but turned down a full interview. One of these moderators stated over direct messages that they felt the tool idea was potentially dangerous for users' well-being because it would create a barrier to connection between support requesters and providers. The other moderator wrote that the tool sounded like it would not improve on the process of using the Reddit search function to find relevant past responses or using a web browser to find online informational resources.

2.5 Data Availability

The dataset we used to train the response quality classifier and the question-answering system, with usernames and post content removed, is available at https://github.com/benji700/qa_project/blob/main/anxiety_subreddit_data_complete.csv. In addition, the corpus of 80 online resources is available at https://github.com/benji700/qa_project/blob/main/MHA%20resources.xlsx.

3 Discussion

In the present work we described our progress in developing a tool to help users rapidly produce high-quality and well-informed responses to anxiety-related questions. Although automated tools exist to help Reddit moderators identify potentially problematic content, to our knowledge, no tools exist to improve response quality to mental health-related questions. Our work demonstrated both opportunities and challenges to designing such a tool.

The resultant tool takes in a text-based question and its associated post metadata from a subreddit and suggests 1) a previous response to a similar question and 2) a relevant online informational resource. We chose to include an online resource due to difficulty in identifying relevant responses to new questions and because many questions expressed recurring themes, so that a relatively small set of online informational resources might be useful to help answer many of those questions. Yet, our results show that both informational and emotional support can be valuable. Online informational resources might be able to provide better informational support, especially when those resources come from high-quality and authoritative sources. Emotional support, however, might need to come from humans who can understand and empathize with one's problems. Users may prefer human-generated responses when seeking empathic support online (Morris et al. 2018), and so efforts to remove a human component from mental health-related QA systems might face limitations.

Moderator perspectives on the tool were mixed. Those with positive impressions of the tool felt it could provide practical support for both moderators and subreddit members interested in rapidly helping posters, helping to offer timely support to more people. Even though they were positive about the tool, these moderators also stressed that if the tool were to be made publicly available for subreddit users, working closely with mental health communities would be essential. Moderators with less favorable impressions of the tool argued that making such a tool available to users would weaken the human connection between support requesters and support providers. They also predicted that the tool would be redundant with web browsers that could be used to look up the relevant information to write a high-quality response. These are useful considerations for other tools seeking to provide support for users on online mental health forums and it would be useful to consider the incremental or differential benefit of such a tool compared to tools like web browsers.

Our QA system and online resource suggestion system outperformed chance, but they did not produce useful suggestions reliably enough to be used without a human in the loop. Despite their inconsistent response quality, these systems might still help human response-providers to suggest higher-quality answers quicker. Even when imperfect, the tool's suggestions may provide useful starting points from which users can then build their own contoured responses. It is worth noting that in our initial hand-coding of responses to anxiety-related questions, about half of the responses were coded as low-quality, so even human-generated responses may not always be useful. Finally, if the tool's ability to produce relevant responses and online resources were to become much more reliable, it could be re-designed to be used by support-seekers themselves to find responses relevant to their own problems.

Next Steps and Future Directions. To improve the QA and online resource suggestion systems, we are designing the tool to produce more accurate recommendations over time based on feedback on response quality from its users. To begin gathering this user feedback, we will release the tool to a group of Reddit moderators and users who will use the tool and rate the quality of its suggestions. Even among this group, careful attention will be necessary to ensure that the tool's recommendations are safe and that users understand its limitations.

In addition to this iterative feedback from users, the suggestion systems could also be improved with a larger corpus of question-response interactions and a more elaborate model. Simplifying the data inputted into the system could also be relevant. For example, it is possible that the questions and responses from anxiety subreddits were too long for the QA system to parse the most relevant part of the question; Morris et al. (2018) designed a similar system using much shorter questions and responses and achieved more reliable response quality.

While the tool has demonstrated some potential, future work will be necessary to know if it can be a valuable contribution to online support communities. Open questions include if the tool produces responses that question-askers find useful, as well as whether people are willing to use the tool. These questions can be answered through experiments, case studies, or by making the tool publicly available and measuring user activity. Further, the tool will need to be refined through user experience research and with input from stakeholders in online mental health communities.

Limitations. First, although we were able to achieve reasonable inter-rater reliability in response quality ratings, our measure of response quality was not a precise estimate of our variable of interest: how much help seekers felt that a given response to their question was useful, informational, and empathic. An alternative strategy would have been to follow Peng et al. (2021) in evaluating help seekers' expressed satisfaction with responses (measured via linguistic analysis of their responses to other users' responses to their posts), although such an approach is also not ideal because many who asked a question do not react to responses they received. Second, our metrics of linguistic features such as sentiment and health-related words were imprecise because they were derived from a simple word-counting computational method and did not account for conversation norms on the subreddits we examined. These metrics would have been more valid with human coding, although such a strategy would have been too labor-intensive to perform on the full dataset and could also introduce bias from raters.

Conclusion. Our tool holds potential for improving the availability of effective support on online anxiety support communities by helping users to quickly produce high-quality responses to anxiety-related questions. As it continues, this project could also provide insights about the utility of crowdsourcing collective wisdom from the internet towards solving challenging personal problems. However, empirical research on our tool's effectiveness is necessary.

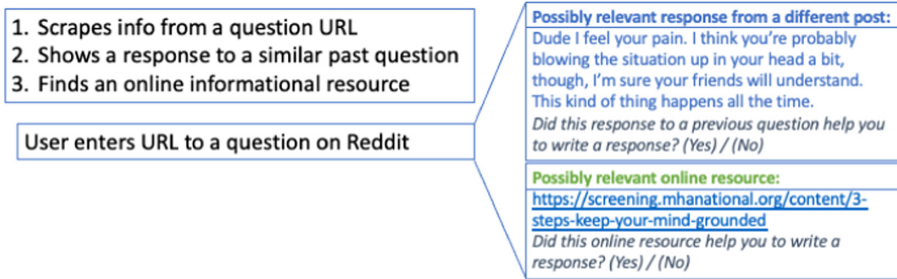


Fig. 3. Design mockup for the response helper tool. Based on moderator feedback, input may instead be automatically scraped from Reddit as the tool user is browsing.

References

- Bargh, J.A., McKenna, K.Y.A., Fitzsimons, G.M.: Can you see the real me? Activation and expression of the “true self” on the Internet. *J. Soc. Issues* **58**(1), 33–48 (2002). <https://doi.org/10.1111/1540-4560.00247>
- Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., Blackburn, J.: The pushshift reddit dataset. In: *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, pp. 830–839 (2020)
- De Choudhury, M., De, S.: Mental health discourse on reddit: self-disclosure, social support, and anonymity. In: *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1 (2014)
- Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D.: Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* **15**(1), 3133–3181 (2014)
- Flesch, R.: A new readability yardstick. *J. Appl. Psychol.* **32**(3), 221 (1948)
- Gkotsis, G., et al.: The language of mental health problems in social media. *CLPpsych@HLT-NAACL* (2016). <https://doi.org/10.18653/v1/W16-0307>
- Kaveladze, B., Schueller, S.: Mental health subreddits offer users valuable and unique support. Poster Session Presented at the 2020 Association for Psychological Science Virtual Poster Showcase (2020). <https://imgur.com/a/RO1wHJd>
- Kavuluru, R., Williams, A.G., Ramos-Morales, M., Haye, L., Holaday, T., Cerel, J.: Classification of helpful comments on online suicide watch forums. In: *ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, pp. 32–40 (2016). <https://doi.org/10.1145/2975167.2975170>
- Kohn, R., Saxena, S., Levav, I., Saraceno, B.: The treatment gap in mental health care. *Bull. World Health Organ.* **82**, 858–866 (2004)
- Mazuz, K., Yom-Tov, E.: Analyzing trends of loneliness through large-scale analysis of social media postings: observational study. *JMIR Ment. Health* **7**(4), e17188 (2020). <https://doi.org/10.2196/17188>
- Morris, R.R., Kouddous, K., Kshirsagar, R., Schueller, S.M.: Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. *J. Med. Internet Res.* **20**(6), e1014 (2018)
- Nielsen, F.Å.: A new ANEW: evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903* (2011)
- O’Leary, K., Bhattacharya, A., Munson, S.A., Wobbrock, J.O., Pratt, W.: Design opportunities for mental health peer support technologies. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 1470–1484 (2017)

- Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
- Peng, Z., Ma, X., Yang, D., Tsang, K.W., Guo, Q.: Effects of support-seekers' community knowledge on their expressed satisfaction with the received comments in mental health communities. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–12 (2021)
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2019). <https://www.R-project.org/>
- Rains, S.A., Tsetsi, E.: Social support and digital inequality: does internet use magnify or mitigate traditional inequities in support availability? *Commun. Monogr.* **84**(1), 54–74 (2017)
- Sharma, E., De Choudhury, M.: Mental health support and its relationship to linguistic accommodation in online communities. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 641:1–641:13 (2018). <https://doi.org/10.1145/3173574.3174215>
- Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* **29**(1), 24–54 (2010)
- Van Rossum, G., Drake, F.L.: *Python 3 Reference Manual*. CreateSpace, Scotts Valley (2009)
- Wainberg, M., Alipanahi, B., Frey, B.J.: Are random forests truly the best classifiers? *J. Mach. Learn. Res.* **17**(1), 3837–3841 (2016)
- Wickham, H., et al.: Welcome to the tidyverse. *J. Open Source Softw.* **4**(43), 1686 (2019). <https://doi.org/10.21105/joss.01686>