# Human-Centered Data Analytics for Identifying Mental Health Risks in Digital Communities

**Srijani Choudhury**

**University of New Haven**

**DOI: https://dx.doi.org/10.51244/IJRSI.2025.1215PH000216**

**Received: 16 November 2025; Accepted: 22 November 2025; Published: 12 December 2025**

## ABSTRACT

We introduce Human-Centered Data Analytics (HCDA), a method for identifying online mental health issues that blends computational techniques with user-centered design principles to produce insights that are practical, interpretable, and morally acceptable. The method entails extracting anonymized text and interactional data from online public spaces, finding characteristics linked to stress, anxiety, and depression, and then using machine learning (ML) in conjunction with user-centered validation techniques to extract first recognition patterns. According to the study's findings, HCDA is more accurate and interpretable than conventional data-driven models at detecting early indicators of mental health disorders. Interpreting risk variables contextually is also informed by user interaction observations. In order to provide proactive, ethical solutions for members of the digital community, the work highlights the value of combining quantitative analysis with human-centered viewpoints. It also has practical implications for mental health experts, platform developers, and policy makers.

**Keywords:** Human-Centered Data Analytics, Mental Health, Digital Communities, Risk Detection, Ethical AI

## INTRODUCTION

The growth of online communities like social media, forums, and online support groups has sped up how people ask for assistance, exchange experiences, and interact. According to De Choudhury et al. (2013), these platforms offer opportunities for social connection, but they also put users at risk for mental health issues like stress, anxiety, and depression. Traditional approaches to mental health monitoring, which are typically restricted to self-reported or clinical assessments, were unable to keep up with the dynamic and context-rich nature of online interactions. 10 Human-Centered Data Analytics (HCDA), which combines computational techniques with human-centered design, offers a possible answer to these problems. When compared to typical data-driven analysis, HCDA is more concerned with interpretability, user context, and ethics, which means that the insights are more than just observations; they identify hazards without infringing on social tensions or privacy (Wang et al., 2021). The HCDA can identify early indicators of mental health problems before they worsen by analyzing behavioral, linguistic, and interactional signals in online communities. This study aims to explore the use of HCDA in identifying mental health risks associated with internet activity. In particular, it examines the integration of human-centered validation and machine learning-based models, and it talks about both the interpretability of the results and the effectiveness of risk identification. A growing need for user-aware and ethical methods to monitor mental health in digital societies is addressed by this work.

## LITERATURE REVIEW

### Digital Communities and Mental Health Support

Online environments Online forums, social media platforms, and virtual support groups are examples of digital communities that have developed into crucial settings for monitoring and offering mental health support. According to research, forums are an online way to give young people information about how to get help and some emotional (and integrated "infomotional") support for mental health concerns (Dewa et al., 2021). Additionally, research indicates that online communities help people become more resilient, especially those living in remote or underdeveloped locations where they may access timely and flexible peer assistance outside of typical healthcare settings (Golz et al., 2022).

Actions in online communities The likelihood of risk being identified using data analytics is increased since user behavior in online communities frequently reflects self-disclosure patterns driven by anonymity and perceived safety (Taylor et al., 2025). Additionally, the user's intended results (self-efficacy and decreased self-stigma) can be significantly impacted by elements like as platform design, moderation, and community climate (Lobban et al., 2025). These findings demonstrate the potential of online communities as abundant sources of information for mental health research. There is a knowledge vacuum in research, nevertheless, as most of the current work focuses on therapeutic benefits rather than predictive risk assessment using human-centric analytics.

**Digital Mental Health and Human-Centred Design**

Human-Centered Design: HCD makes sure that any solution is human, intelligible, accessible, and ethically developed by concentrating on the people who will eventually benefit from a technology. Despite the extremely technical nature of AI and analytics, HCD is still underutilized in the field of digital mental health (Haroun et al., 2022). Sensor-rich and AI-based systems predominated in an assessment of data-driven health ecosystems, however the difficulty of integrating humans and machines such as interpretability and, most crucially, trust—is prominently highlighted (Hummel et al., 2024).

Digital communities must adhere to human-centered design principles in order to guarantee that analytics models are actionable and comprehensible. They aid in integrating findings into suitable solutions that respect ethics, consent, and privacy. Data-driven models and real user contexts can be reconciled by HCDA through participatory design and co-creation with users (Tandon, 2024).

**Artificial Intelligence and Psychology in Digital Mental Health**

Social media and online communities have made extensive use of machine learning, natural language processing, and big data analytics to detect mental health problems. According to a narrative analysis of 25 years of research, computer models can forecast the likelihood of self-harm, anxiety, and depressive symptoms based on posting behaviors, language patterns, and even social network structure (Tandon et al., 2024). Similarly, text mining research conducted in online communities during the COVID-19 pandemic revealed that users' communication may reveal their coping mechanisms and mental suffering (Golz et al., 2022).

Additionally, there is solid proof that community organization and moderation have a significant impact on the caliber and reliability of analytics data. Because user interactions in well-moderated forums are more structured and context-specific, they not only encourage participation but also improve the interpretability of machine learning models (Lobban et al., 2025). However, the majority of these models have problems with interpretability, user adoption, longitudinal model tracking, and ethical implementation.

**HCDA Integration in Digital Communities: Research Gaps and Synthesis**

Opportunities and challenges for HCDA are shown by combining insights from analytics, user-centered design, and digital community behavior:

Human-Centered Analytics in Community Contexts: Word-Vectors 23 Although there is a dearth of research on the coupling of predictive models with human-centered design for digital peer assistance contexts, some hopeful findings based on similar themes are presented.

Interpretability and Actionability: According to Hummel et al. (2024), a lot of analytics models might act as "black boxes" that make it difficult for moderators, practitioners, or users to apply the results.

Privacy and Ethics: Mining user-generated information raises ethical questions concerning consent, anonymity, and possible algorithmic prejudice. Analytical pipelines must to contain ethical frameworks (Haroun et al., 2022).

Language, cultural norms, and online platform affordances distinguish the environment of the digital community from that of the clinic. It is necessary to assess these risk factors in light of the previously mentioned scenario (Taylor et al., 2025).

Evaluation and Longitudinal Impacts: The majority of research is descriptive or cross-sectional. The idea that analytical treatments lead to real improvements in mental health outcomes requires long-term data (Bevan Jones et al., 2022).

This research demonstrates that while digital communities offer chances to identify mental health hazards, there is still much to learn about integrating human-in-the-loop methods with sophisticated analytics. This area highlights the value of HCDA in creating morally sound, understandable, and practical solutions for digital mental health environments.

## METHODOLOGY

### Design and Environment of the Study

A mixed-method human-centered data analytics (HCDA) methodology was employed in this study to identify risk factors for mental health in online communities. The project combined (1) qualitative investigation of human situations, (2) computation sub-segment data collection and analytics, and (3) iterative validation with community stakeholders and mental health professionals using a human-centered design methodology (Vial et al., 2022).

Participants and the Context of the Community The study focused on users of digital communities, particularly social media groups and online forums for low-income urban populations. Purposive sampling was used to passively recruit participants for qualitative interviews (n = 20–30) who were community moderators, regular contributors, and mental health support facilitators. At the same time, in accordance with ethical standards, unidentifiable digital interaction data (posts, comments, and metadata) were recorded on public online platforms (Butorac et al., 2025).

### Gathering and Preparing Data

Text content, temporal metadata, and interaction metrics were all included in the digital data. Among the pre-processing procedures were the creation of engagement metrics, anonymization, and text standardization and tokenization (Owen et al., 2024). In accordance with the human-centered phase, feature engineering included language and behavior-adjusted patterns.

Extracting Features and Creating Models Features were divided into the following categories: Language-based: subject modeling, emotional tone, pronoun usage, and sentiment Behavioral/Temporal: posting frequency, periods of dormancy

Interactional: quantity of responses, thread exchange, and social network centrality These characteristics were supplied as input to machine learning classifiers that predict mental health risk, such as random forest and gradient boosting. Accuracy, precision, recall, and interpretability metrics, such as feature significance and SHAP values, were used to evaluate the models' performance (Tandon et al., 2024).

Validation and input from stakeholders Stakeholders (a panel of moderators and mental health specialists) evaluated the model's results for interpretability and usefulness. According to Hummel, Braun, and Bischoff (2024), workshops have been used to extend features, check flagged cases, and involve ethical issues.

Moral Aspects to Take into Account Ethics clearance was acquired. All data was de-identified, and consensus from the community was gathered. The focus was on algorithmic fairness, privacy, and transparency (Butorac Mathieu et al., 2025).

# RESULTS

### Qualitative Results

Interviews and workshops yield various important insights:

Data and confidence shared with others: Respondents emphasized having control over analytics and a transparent usage of data (Butorac et al., 2025). Cultural and linguistic context: euphemism terms, code-switching, and local idioms were essential in recognizing distress indicators (Taylor et al., 2025).

Changes in behavior as a precursor to rising mental risk: For instance, in several instances, the abrupt stop of postings stated above, followed by intense emotional outpouring, signaled a crisis (Lobban et al., 2025a).

Stigma and barriers to getting help: Respondents referred to broad expressions of distress rather than the labels "depression" or "anxiety" (Dewa et al., 2021).

Alerts must be interpretable: Before taking action, stakeholders requested clear explanations of the case they were informed about (Hummel et al.

### Quantitative Results

with an AUC of 0.82, accuracy of 0.75, and recall of approximately 0.70, gradient boosting was able to distinguish between high-risk and low-risk users.

Changes in thread involvement, spikes in negative emotion, and gradual declines in activity were the main characteristics.
Prediction was much enhanced by linguistic change indicators, such as a greater use of first-person singular pronouns and a decrease in positive emotions.

Despite a 15% false positive rate, 60% of model-flagged users were confirmed to have risky behavior recognized by moderators by stakeholder assessment. Integration of Human-Centered Design and Analytics

Stakeholder trust and feature relevance were enhanced by HCDA:

Importance of the feature: The discovery of more precise risk indicators was made possible by cultural and community intelligence.

Acceptability and intelligence: The stakeholders' active participation raised the prediction system's utilization and level of confidence (Tandon, 2024).

### Limitations:

Since some human eyes must examine the data, false negatives continue to be a problem. The sparsity issue is introduced by low-participation users. Opt-in consent and ongoing oversight are required due to ethical considerations. The estimation of long-term forecasts is limited by the short validation period.

# DISCUSSION

In this study, we examine the potential applications of Human-Centered Data Analytics (HCDA) methodologies to identify mental health hazards in online communities by fusing qualitative user-centered insights with computational approaches. The findings highlight the potential and risks of HCDA in proactive mental health monitoring.

Human-centered design and analytics using a blended model The results underscore the significance of machine learning methodologies at the nexus of human-centered design. According to qualitative research, community aesthetic norms, cultural metaphors, and regional idioms are crucial for deciphering online behavior. Models can identify minor signs of discomfort that are not yet detectable by simply data-driven methods by including these insights into feature engineering (Vial et al., 2022; Haroun et al., 2022). This is consistent with earlier research that demonstrated how human-centered design can enhance DMDIs' interpretability and applicability (Tandon, 2024).

Behavioral Insights and Predictive Features Quantitative analysis revealed that sentiment alterations, interactional changes, and temporal posting patterns were the strongest predictive indicators of mental health risk. These findings are in line with earlier studies that showed a decrease in active participation, a rise in the usage of words associated with negative emotions, and alterations in response patterns that took place before visible signs of distress (Owen et al., 2024; Lobban, Caton & Glossop, 2025). By incorporating human-centered validation, our model became helpful to moderators and practitioners because these features were significant in the community setting.

### Implications for Interventions in Digital Communities

Predictive analytics can complement conventional mental health promotion in online communities, as demonstrated by the HCDA paradigm. Moderators of community messages and mental health specialists can take preemptive steps to offer assistance or make an ethical intervention by using interpretable signals that are driven by the features' context. It emphasizes that risk sensing in online communities is a problem that requires consideration of ethical, privacy, and trust issues rather than being solely a technical one (Butorac et al., 2025; Hummel, Braun & Bischoff, 2024).

### Taking Care of Limitations

Nevertheless, despite the positive outcomes, certain limits were discovered. False negatives highlight the value of human control, particularly for users with little engagement and sparse data. Even though ARIMA has a great long-term forecasting capacity, the short validation time further limits the extent. Future research might look at continuous engagement metrics and longitudinal models, as well as create multi-platform data fusion to improve information dependability. Another strategy to resolve privacy concerns and boost overt trust is to implement opt-in features and real-time feedback loops (Taylor, D'Alfonso, & Dolan 2025).

### Theory and Practice Contribution

By demonstrating how HCDA can act as a mediator between technological analytical and community-respecting mental health interventions, this research contributes to the body of literature. Even if we assume the popularity contest, it shows the value of human-controlled insights in an effective model and how, when prediction is in line with lived experiences in digital communities, stakeholders are more likely to trust it. For vulnerable populations in the urban low South (such as Dhaka's slum neighborhoods), the findings also suggest a methodology for incorporating morally and culturally appropriate analytical solutions into human-centered design techniques.

### Prospects for the Future

Future research could expand HCDA to include cross-cultural settings, real-time intervention models, and longitudinal tracking. Early identification and intervention could be enhanced by integrating with digital medicines or AI-based recommender systems. Additionally, incorporating end users into model creation and evaluation through more participatory ways might improve trust, reduce algorithmic bias towards staff, and encourage adoption in community contexts.

## CONCLUSION

We demonstrate how Human-Centered Data Analytics (HCDA), which combines computing and human-derived ideas from online communities, best catches online mental health danger signals. The patterns demonstrate that, when placed inside a human-centric design framework, linguistic patterns, temporal posting behaviors, and interactional metrics can be accurately predictive of mental health risk-tendency. Predictive models became more relevant and interpretable when human-centered insights were incorporated, making the results useful and practical for mental health professionals and community moderators (Vial, Boudhraâ, & Dumont, 2022; Haroun, Sambaiga, & Sarkar). Furthermore, involving stakeholders in the process validation improved the analytical framework's practical usefulness, ethical sensitivity, and trust (Tandon 2024; Butorac et al., 2025). The need to further improve and prospectively evaluate the use of such sensors in a longitudinal fashion with suitable ethical precautions is highlighted by the false negative rates, lack of data for low-activity users, and validation over brief periods. In order to guarantee that analytics tools are culturally appropriate and contextually relevant, future study should focus on long-term monitoring, real-time intervention techniques, cross-cultural generalizability, and further enhance participatory research methodology. By showing that HCDA is a workable, morally acceptable, and interpretable method for early identification of mental health hazards from online communities, especially among underrepresented or vulnerable populations, this study advances the theory and practice of digital mental health. From the standpoint of creating technically sound digital mental health solutions that satisfy users' demands, cultural context, and ethical standards, the ramifications of these findings are significant.

## REFERENCES

1. Butorac, I., McNaney, R., Seguin, J. P., Northam, J. C., Tully, L. A., Carl, T., & Carter, A. (2025). Developing digital mental health tools with culturally diverse parents and young people: Qualitative user-centered design study. *JMIR Pediatrics and Parenting, 8*, e65163.
2. Dewa, L. H., Lawrance, E., Roberts, L., Brooks-Hall, E., Ashrafian, H., Fontana, G., & Aylin, P. (2021). Quality social connection as an active ingredient in digital interventions for young people with depression and anxiety: Systematic scoping review and meta-analysis. *Journal of Medical Internet Research, 23*(12), e26584.
3. Haroun, Y., Sambaiga, R., & Sarkar, N. (2022). A human-centered approach to digital technologies in health care delivery among mothers, children, and adolescents. *BMC Health Services Research, 22*, 1393.
4. Hummel, P., Braun, M., & Bischoff, S. (2024). Perspectives of patients and clinicians on big data and AI in health: A comparative empirical investigation. *AI & Society, 39*, 2973–2987.
5. Lobban, F., Caton, N., & Glossop, Z. (2025). Impacts of using peer online forums in mental health: Realist evaluation using mixed methods. *Journal of Medical Internet Research*.
6. Owen, D., et al. (2024). AI for analyzing mental health disorders among social media users: Quarter-century narrative review of progress and challenges. *Journal of Medical Internet Research, 26*, e59225.
7. Tandon, A., Cobb, B., & Centra, J. (2024). Human factors, human-centered design, and usability of sensor-based digital health technologies: Scoping review. *Journal of Medical Internet Research, 26*, e57628.
8. Taylor, T., D’Alfonso, S., & Dolan, M. J. T. (2025). How do users of a mental health app conceptualise digital therapeutic alliance? A qualitative study using the framework approach. *BMC Public Health, 25*, 2450.
9. Vial, S., Boudhraâ, S., & Dumont, M. (2022). Human-centered design approaches in digital mental health interventions: Exploratory mapping review. *JMIR Mental Health, 9*(6), e35591.