CROWDSOURCED MENTAL HEALTH DATA PLATFORM

Problem Identification Report

Project Group 12

Hongwu Li (40280054),
Jananee Aruboribaran (40129224),
Hanieh Maleki (40217050),
Priscilia Momo (40155196),
Yu-Hang Lin (40266597)

Table of Contents

Problem Identification Report

I. Problem/Opportunity Statement

A. The Mental Health Data Crisis: Gaps, Challenges, and Opportunities

While mental health disorders affect **one billion people worldwide** [1], barriers in data collection systems and widespread stigma along with fragmented healthcare frameworks obstruct both effective treatment and comprehensive research. Approximately **one-third of Canadians** [2] will face mental health problems during their lifetime yet marginalized groups endure greater obstacles from inadequate data collection and societal stigma with Indigenous communities showing 15% greater psychological distress than national levels [3] and 2SLGBTQI+ youth having twice the suicide attempt risk compared to their peers [4]. Real-time mental health trends remain undetected through traditional assessment methods which depend on **psychometric surveys** and clinical checklists because these methods **exclude vulnerable populations** including homeless people and individuals in hospitals or correctional facilities. **The methodological restrictions** present in these approaches frequently conceal the actual variety of mental health experiences.

While **three-quarters of people** in low- and middle-income countries (LMICs) do not receive mental health care access, but high-income countries face fragmented data systems and stringent privacy laws such as GDPR and PIPEDA. **Mobile applications and wearable technology** offer potential solutions for ongoing **personalized tracking** of mental health measures like patterns in sleep and mood changes. **Data silos** like proprietary app databases which fail to connect with clinical systems along with unresolved privacy issues limit their potential. Research from 2019 demonstrated that half of the leading depression and smoking cessation apps transmitted user data to external entities without **obtaining users' explicit permission** [6], which damaged user trust and reduced app engagement.

These challenges stem from multiple **root causes:** Multiple root causes generate these challenges including stigma and privacy fears that prevent people from sharing mental health issues alongside fragmented systems that hinder data interoperability combined with **research biases** excluding important voices from clinical studies. **Delayed insights** lead to ineffective interventions while poor or fair perceived mental health among Canadian youth saw a doubling rate between 2011 and 2018 [7]. Experts recommend the implementation of **enhanced data sharing** through the **FAIR** framework to improve worldwide mental health research results and ensure **equitable outcomes.** [8]

There exists a distinct opportunity to establish a **digital platform** which allows people to **confidentially submit** mental health experiences thereby collecting crowdsourced data. A digital platform would reach a worldwide audience since **68% of the global population** [9]

has internet access. **Protecting privacy** and **diminishing stigma** will drive wider participation which will generate **more comprehensive datasets** that remedy current issues with data diversity and access. The **real-time analysis** of combined user data has the potential to revolutionize mental health research by transforming it from a reactive practice to proactive and adaptable research methods which enable researchers and other stakeholders to better **identify trends** and optimize treatment strategies for enhanced **global** mental health outcomes.

B.   Why Mental Health Data Innovation Matters

Resolving systemic obstacles plays a crucial role in propelling mental health research forward which leads to better care delivery. The annual financial toll of untreated mental illness amounts to **$50 billion** in lost productivity and healthcare costs in Canada while depression and anxiety generate a global economic burden surpassing **$1 trillion** (MHCC, 2017; The Lancet, 2020 [10]). The combination of stigma and inadequate culturally sensitive services raises the obstacles to healthcare for marginalized communities and intensifies existing disparities. The COVID-19 pandemic highlighted the critical need for timely data when mental health demands surged but service providers did not have access to precise real-time information for resource allocation. [11]

A **crowdsourced mental health data platform** that captures real-time information can deliver transformative benefits across several domains:

- **Research**: Extensive and diverse datasets can speed up discovery while shortening the traditional **17-year delay** [12] between research breakthroughs and clinical application.
- **Policy**: Evidence-based resource allocation remains essential for remote or underserved regions like Canada's rural North which often face limited mental health services.
- **Healthcare**: Timely insights in healthcare reduce severe psychiatric misdiagnosis rates which currently stand at 40% and allow providers to offer more personalized patient care. [13]
- **Global Collaboration**: **International research and policy partnerships** will strengthen mental health initiatives through unified data standards and joint knowledge sharing.

Without exploiting these opportunities, the world economy could endure **$6 trillion losses** by 2030 [14] while deepening social and health disparities as projected by the World Economic Forum's 2025 report. Researchers gain unprecedented depth and scope of understanding when they use **modern crowdsourced platforms** to collect mental health data which enhances both data **quality and inclusivity**. This approach enables researchers to create targeted evidence-based strategies through **real-time diverse** input which leads to more equitable and effective mental health outcomes globally.

II. Stakeholder Analysis

A. Identification of Key Stakeholders

A mental health data platform that relies on user-contributed information operates as a common space where stakeholders from diverse backgrounds unite. The various stakeholders possess individual interests along with their own potential benefits and concerns which require careful management. The diagram below displays the stakeholders we have identified.

For each of the main categories of stakeholders identified in the mind map, we provide below an overview of the value this project could provide them:

1.   Table 1: Stakeholder Overview and Potential Benefits

Stakeholder Group	Description	Potential Benefits
End Users (Individuals with Mental Health Experiences, Caregivers, and Families)	People with lived mental health experiences; their families and caregivers may also engage.	- Access to collective mental health insights. - Greater representation in research. - Reduction in stigma through shared experiences.
Researchers (Academic & Clinical Researchers, Data Scientists, Public Health Experts)	Experts studying mental health trends, treatment efficacy, and population health.	- Access to a large, diverse, real-time dataset. - Accelerated mental health discoveries.
Healthcare Providers (Psychiatrists, Psychologists, Therapists, General Practitioners, Social Workers)	Professionals treating mental health conditions who could use insights from the platform.	- Improved diagnosis and treatment strategies .- Earlier detection of mental health crises.
Community Organizations & NGOs (Mental Health Advocacy Groups, Crisis Helplines, Support Networks)	Groups working to support mental health awareness, crisis intervention, and advocacy.	- Data-driven advocacy for better mental health policies. - More effective crisis response.
Policymakers & Regulatory Bodies (Government Health Departments, Ethics Committees, Data Protection Agencies, International Health Organizations)	Authorities responsible for mental health policies, regulations, and legal compliance.	- Data-backed policymaking for better resource allocation. - Identification of underserved communities.
Potential Sponsors / Funding Bodies (Private Sector, Philanthropic Organizations, Research Grants, Government Agencies)	Organizations that may provide financial, technological, or infrastructural support.	- Social impact recognition. - Advancement of mental health initiatives and research.

Project Manager	Oversees project planning, coordination, and execution.	- Ensures smooth development and timely delivery of the platform. - Manages stakeholder expectations and project goals.
Technical Team (Developers, UX/UI Designers, Data Engineers, Cybersecurity Experts, Legal Advisors)	Responsible for designing, developing, securing, and maintaining the platform.	- Creates a secure, scalable, and user-friendly platform. - Ensures compliance with data protection laws.

In the following table, for each stakeholder, we outline their role in the project, along with their key interests and primary concerns.

2.   Table 2: Stakeholder Roles, Interests, and Challenges

Purpose: Detail each stakeholder's role in the project, their key interests, and concerns.

Stakeholder Group	Role	Key Interests	Primary Concerns / Challenges
End Users	Share anonymous mental health experiences on the platform.	Anonymity, Support, Community, User-friendly platform	Privacy risks, Data security, Fear of stigma
Researchers	Use crowdsourced data for studies and identifying mental health trends.	Access to structured and diverse dataset, Up-to-date data	Data accuracy, Sampling bias, Ethical research use
Healthcare Providers	Leverage insights to refine treatment plans and interventions.	Reliable data, Patient insights, Early intervention, Evidence-based care	Integration with clinical systems, HIPAA/GDPR compliance, Misinterpretation of data
Community Organizations & NGOs	Utilize insights for advocacy and service improvement.	Community needs assessment, Targeted interventions, Resource allocation	Underrepresentation of marginalized groups, Public skepticism, Effective dissemination of insights
Policymakers & Regulatory Bodies	Use aggregated insights for policy decisions.	Evidence-based policymaking, Public health monitoring	Data privacy laws, Policy feasibility, Ensuring unbiased reporting
Potential Sponsors / Funding Bodies	Provide financial and infrastructure support for the platform.	Social impact, Innovation, ROI	Ethical concerns, ROI justification, Long-term viability

Project Manager	Oversees project development and team coordination.	Project success, Stakeholder alignment, Risk management	Meeting deadlines, Resource allocation, Balancing needs
Technical Team	Build, maintain, and secure the platform.	Secure architecture, Scalable infrastructure, Compliance with regulations	Integration challenges, Compliance, Cybersecurity risks

III. Relevance to the Software Solution

A. Addressing the Problem Through Software Development

The crowdsourced mental health data platform minimizes the limitations associated with traditional mental health data collection by utilizing cutting-edge digital technology. Using blockchain, secure cloud storage, decentralized identity management, and AI-driven data processing, the solution software will guarantee the following:

1. Ensure Anonymity and Privacy:

Problem: Traditional mental health platforms require users to create an account; this is fraught with potential issues related to privacy breaches, data breach, and disclosure stigma associated with mental health disclosures.

Solution: Anonymous Mental Health Data Submission Using Decentralized Identifiers (DID)
- Users create a Decentralized Identifier (DID) for submitting experiences. DID is a cryptographically secure, non-traceable identity that allows the user to submit, update, or delete entries without personally identifiable information linking. Self-Sovereign Identity (SSI) ensures users have full control over their data.

2. Enhance Data Integrity and Transparency:

Problem: Mental health can be fudged or otherwise manipulated, thus making it difficult for researchers and policymakers to trust trends derived from user submissions.

Solution: Blockchain for Data Integrity and Transparency
- A Blockchain ledger records metadata like timestamp, hash of submissions, to make the experiences tamper-proof and verifiable. Smart contracts control access, thus allowing users to give/take away consent on the use of anonymized data in research studies.

3. Preserve Data Security During Processing:

Problem: Traditional cloud storage methods necessitate decryption before processing, risking leakage and unauthorized access to sensitive information related to mental health.

Solution: Secure Cloud & Data Storage for Privacy-Preserving Insights

- The platform employs secure cloud storage techniques, such as homomorphic encryption. Examples include Microsoft SEAL and IBM FHE, which will enable AI and analytics tools to compute insights directly on encrypted data without decryption. Multi-Party Computation allows different researchers to analyze encrypted data sets in collaboration, without seeing any user-submitted data. Even platform administrators cannot access raw user data, ensuring zero-knowledge data processing.

4.   Automate Data Structuring and Anonymization:

Problem: It is inefficient to manually process such large volumes of data on mental health experiences, and traditional methods cannot efficiently anonymize access to the same.

Solution: AI-Driven Data Structuring and Anonymization
- Unstructured text is transformed into structured, searchable datasets. Classify user experiences with labels naming the mental health condition-for example, anxiety, depression, PTSD. Extract insights without linking data back to individual users.

In addition to providing researchers with high-quality, trustworthy data, these integrated technologies solve the fundamental issues in a way that gives users the confidence to talk about their experiences.

B.   Scope of the Software Solution

1.   Features & Functionalities:

- **Anonymous Data Submission:** While generating, modifying, or removing their mental health experiences, users will have complete ownership over their data and be able to generate and use Decentralized Identifiers (DIDs).

- **Blockchain for Data Integrity:** Information such as timestamps and hashes that provide authenticity, transparency, and user-controlled consent management with regard to research access are stored in a tamper-evident ledger.

- **Privacy-Preserving Data Storage & Processing:** Using multi-party computation, homomorphic encryption, and secure cloud storage to gain insights without requiring access to user raw data.

- **AI-Driven Data Structuring & Anonymization:** In order to conduct significant insight without violating confidentiality, AI should classify unstructured text for structured data.

- **User-Controlled Consent Management:** Users would have complete control over whether to allow a researcher access to their anonymized personal data.
- **Researcher Access Portal:** Verified researchers will be able to find trends and patterns in mental health experiences by analyzing anonymized and encrypted datasets.

2.   Technological Constraints

- **Decentralized Identity Management:** Blockchain technology and cryptography standards will serve as the foundation for DID and SSI.
- **Blockchain & Smart Contracts**: Blockchain technology, such as Hyperledger or Ethereum, should strike the best possible balance between cost, efficiency, and transparency.
- **Homomorphic Encryption & MPC:** Computationally intensive encryption techniques may require optimization for real-time analysis.
- **AI Processing & Data Structuring:** Requires high computational power and efficient models for text classification and anonymization.

3.   Limitations & Exclusions

- The platform will not provide any real-time mental health consultation or intervention.
- It will not store or process personally identifiable information of any kind.
- It will not support direct user-to-user communication in order to avoid potential misuse.
- It does not replace traditional clinical diagnosis or therapy.

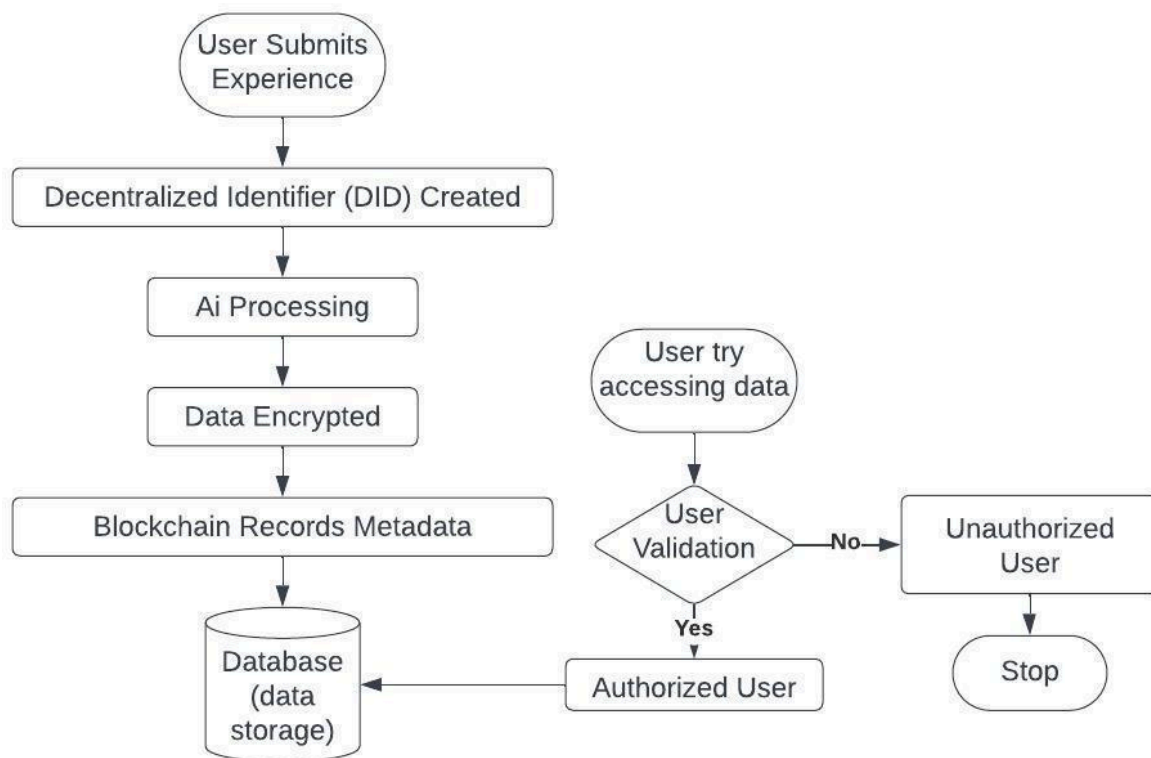## C.   User Data Submission & Access Flow

Figure 1: User Data Flow

1. User Submits Experience
   The user submits their experience or data.
2. Decentralized Identifier (DID) Created
   A decentralized identifier is generated for the submitted data.
3. AI Processing
   The data undergoes processing through AI.
4. Data Encrypted
   The processed data is encrypted for security.
5. Blockchain Records Metadata
   Meta-data of the encrypted data is stored on the blockchain.
6. Database (Data Storage)
   The actual encrypted data is stored in a database.
7. User Tries Accessing Data
   A user attempts to access the data.
8. User Validation
   The system checks whether the user is authorized:
     a. If **yes**, the user is classified as an Authorized User and can access the data through an API.
     b. If **no**, the user is classified as an Unauthorized User and the process stops.

D.   Compliance with privacy laws

The proposed platform is designed to adhere strictly to data privacy regulations such as GDPR, PIPEDA, and HIPAA, ensuring user data is handled with the utmost care and transparency.

a. **Data Minimization:** Collects only essential information necessary for research purposes, aligning with GDPR's principle of limiting data collection to what is directly relevant.
b. **User-Controlled Consent:** Incorporates clear consent mechanisms, allowing users to understand and manage how their anonymized data is utilized. This approach complies with PIPEDA's emphasis on obtaining meaningful consent.
c. **Encryption and Security:** Employs advanced encryption techniques to protect data during transmission and storage, meeting HIPAA's requirements for safeguarding protected health information.
d. **Anonymization:** Utilizes decentralized identifiers and self-sovereign identity frameworks to ensure that personal identities are not linked to the submitted data, further protecting user privacy.

E.   Future Scalability and Integration

The platform is designed with scalability in mind, allowing for future enhancements and broader applications:

- **Integration with Health Apps and Wearables:** Plans to interface with devices like Apple Health and Fitbit to enrich data collection with physiological metrics, providing a more comprehensive view of mental health indicators.
- **Expansion to Other Healthcare Sectors:** The underlying framework can be adapted to collect and analyze data related to chronic illnesses, facilitating research and interventions across various health domains.
- **Federated Learning Implementation:** Exploring federated learning approaches to improve AI models collaboratively without the need for centralized data storage, thereby enhancing privacy and reducing the risk of data breaches.

By leveraging blockchain, AI, and decentralized identity systems, we can transform mental health data collection into a secure, scalable, and inclusive global process.

Sources

[1] "Who Highlights Urgent Need to Transform Mental Health and Mental Health Care." *World Health Organization*, 17 June 2022, www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care#:~:text=In%202019%2C%20nearly%20a%20billion,6%20years%20lived%20with%20disability.

[2] "Data Blog." *Mental Illness in Canada - Data Blog - Public Health Infobase | Public Health Agency of Canada*, 8 Oct. 2020, health-infobase.canada.ca/datalab/mental-illness-blog.htm.

[3] Anderson, Thomas. "Portrait of Youth in Canada: Data Report." *Chapter 4: Indigenous Youth in Canada*, Government of Canada, Statistics Canada, 1 Dec. 2021, www150.statcan.gc.ca/n1/pub/42-28-0001/2021001/article/00004-eng.htm?utm_source=chatgpt.com.

[4] Canada, Women and Gender Equality. "Government of Canada." *Canada.Ca*, / Gouvernement Du Canada, 31 May 2024, www.canada.ca/en/women-gender-equality/free-to-be-me/federal-action-2slgbtqi-communities/facts-stats.html?utm_source=chatgpt.com.

[5] Morris, Kim. "A Path for Scaling Mental Health Treatment in Low-Income Countries: Results from the Sharp Trial." *Institute for Global Health and Infectious Diseases*, 17 Sept. 2024, globalhealth.unc.edu/2024/04/a-path-for-scaling-mental-health-treatment-in-low-income-countries-results-from-the-sharp-trial/?utm_source=chatgpt.com.

[6] K. Huckvale, et al. "Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation." *Semantic Scholar*, JAMA Network Open, 1 Apr. 2019, www.semanticscholar.org/paper/Assessment-of-the-Data-Sharing-and-Privacy-of-Apps-Huckvale-Torous/8ae1d78fdf31ea56aff07b0a5fd830a277e397c8.

[7] Wiens, K, et al. "A Growing Need for Youth Mental Health Services in Canada: Examining Trends in Youth Mental Health from 2011 to 2018." *Epidemiology and Psychiatric Sciences*, U.S. National Library of Medicine, 17 Apr. 2020, pmc.ncbi.nlm.nih.gov/articles/PMC7214527/.

[8] Sadeh, Yaara, et al. "Opportunities for Improving Data Sharing and Fair Data Practices to Advance Global Mental Health: Cambridge Prisms: Global Mental Health." *Cambridge Core*, Cambridge University Press, 3 Mar. 2023, www.cambridge.org/core/journals/global-mental-health/article/opportunities-for-improving-data-sharing-and-fair-data-practices-to-advance-global-mental-health/EAF84DE4FF9300E7CC1213ED0D24C857.

[9] "Digital around the World - Datareportal – Global Digital Insights." *DataReportal*, datareportal.com/global-digital-overview.

[10] "Strengthening the Case for Investing in Canada's ..." *Mental Health Commission*, Mar. 2017, www.mentalhealthcommission.ca/wp-content/uploads/drupal/2017-03/case_for_investment_eng.pdf.

[11] Duden, Gesa Solveig, et al. "Global Impact of the COVID-19 Pandemic on Mental Health Services: A Systematic Review." *Journal of Psychiatric Research*, U.S. National Library of Medicine, Oct. 2022, pmc.ncbi.nlm.nih.gov/articles/PMC9392550/?utm_source=chatgpt.com.

[12] R;, Rubin. "It Takes an Average of 17 Years for Evidence to Change Practice-the Burgeoning Field of Implementation Science Seeks to Speed Things Up." *Semantic Scholar*, U.S. National Library of Medicine, 5 Apr. 2023, pubmed.ncbi.nlm.nih.gov/37018006/.

[13] Ayano G;Demelash S;Yohannes Z;Haile K;Tulu M;Assefa D;Tesfaye A;Haile K;Solomon M;Chaka A;Tsegay L; "Misdiagnosis, Detection Rate, and Associated Factors of Severe Psychiatric Disorders in Specialized Psychiatry Centers in Ethiopia." *Annals of General Psychiatry*, U.S. National Library of Medicine, 2 Feb. 2021, pubmed.ncbi.nlm.nih.gov/33531016/.

[14] Forum, World Economic. "Strategic Intelligence: World Economic Forum." *Stategic Intelligence*, intelligence.weforum.org/topics/a1Gb0000000pTDbEAM.