NATIONAL RESEARCH PROGRAM
UNLOCKING THE MIND: ADVANCING NEURO AND MENTAL HEALTH RESEARCH CALL

<table>
<tr><td>ANMR ID</td><td colspan="3">ANMR01-0217-250042</td></tr>
<tr><td>Project Title in English</td><td colspan="3">Sakina: AI-Powered Culturally-Aware ChatBot for Mental Health Providers Augmentation</td></tr>
<tr><td>Project Title in Arabic (optional)</td><td colspan="3"></td></tr>
<tr><td>Type of Application</td><td colspan="3">◉ New</td></tr>
<tr><td>Type of Research</td><td colspan="3">◉ Basic Research
○ Applied Research
○ Experimental Development / Translational Research</td></tr>
<tr><td>Submitting Institution</td><td colspan="3">Hamad Bin Khalifa University (HBKU)</td></tr>
<tr><td>Lead PI (title, name, position)</td><td colspan="3">Dr. Safa Messaoud, AI Scientist, Qatar Computing Research Institute, HBKU</td></tr>
<tr><td>List of Participants PIs, collaborative institutions, PI residency)</td><td colspan="3">
<table>
<tr><th>PI Name</th><th>Position</th><th>PI Institution</th><th>PI Residency</th></tr>
<tr><td>Dr. Leena Babiker (PI)</td><td>Psychiatry Resident and Founder of Sakina</td><td>QSTP</td><td>Qatar</td></tr>
<tr><td>Dr. Ines Arous (PI)</td><td>Assistant Professor of Computer Science</td><td>York University</td><td>Canada</td></tr>
<tr><td>Dr. Zainab Imam (PI)</td><td>Psychiatrist, Division Chief, Women's Mental Health</td><td>Sidra Medicine</td><td>Qatar</td></tr>
<tr><td>Dr. Hooman Keshavarzi (Consultant)</td><td>Program director for the Masters in Counseling Islamic Psychology</td><td>HBKU</td><td>Qatar</td></tr>
</table>
</td></tr>
<tr><td>Total Funding Requested</td><td>USD 279,550</td><td>Project Duration</td><td>24 months</td></tr>
</table>

Table of Contents

0. PROPOSAL SUMMARY 3
1. INTRODUCTION 3
1.1. BACKGROUND AND PRELIMINARY DATA 3
1.2. RESEARCH HYPOTHESIS AND OBJECTIVES 5
1.3. EXPECTED OUTCOME 6
2. IMPORTANCE TO QATAR 6
2.1. ALIGNMENT TO CALL'S RESEARCH PRIORITY AREAS AND RELEVANCE TO THE LOCAL CONTEXT 6
2.2. SOCIAL, HEALTH, ECONOMIC, AND ENVIRONMENTAL IMPACT 6
2.3. COLLABORATION AND INVOLVEMENT OF END-USERS 7
2.4. UTILISATION OF THE JOINT-FUNDER RESOURCES (Not APPLICABLE) 7
3. WORK PLAN 7
3.1. METHODOLOGY 7
3.2. TECHNICAL DESCRIPTION BY WORK PACKAGE 7
3.2.1. Work Package (WP 1) #: Culturally Aware Library/Knowledge graph for Women Mental Health 7
3.2.2. Work Package (WP 2) #: Offline Augmented Care Plan Generation 8
3.2.3. Work Package (WP 3) #: Offline Care Plan Execution 10
3.2.4. Work Package (WP 4) #: Evaluation Metrics 11;,
3.3. ESTIMATED TIMELINE 12
4. PROJECT MANAGEMENT AND SYNERGY: 12
4.1. SYNERGY AND RELEVANCE OF THE RESEARCH TEAM: 12
4.2. ETHICAL AND REGULATORY REQUIREMENTS: 12
4.3. DISSEMINATION AND EXPLOITATION OF RESULTS: 12
5. REFERENCES 13
6. APPENDIX 17

Research Plan

0. PROPOSAL SUMMARY

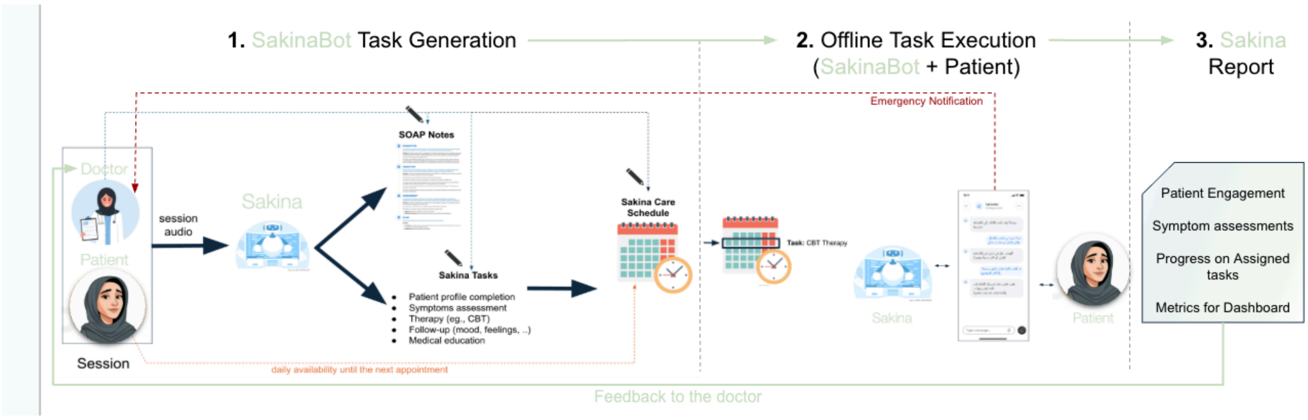

Fig1. Sakina overview

Goal: Sakina (Fig.1) is **the first chatbot designed to assist psychologists and psychiatrists in their practice**. Unlike existing mental health support tools, Sakina operates within a clinical framework, working alongside doctors to **enhance diagnosis and therapy**. Its core mission is to handle tasks assigned by psychologists and psychiatrists **offline between sessions with the mental health provider**, ensuring continuous care and improving the quality of treatment. As an initial focus, **Sakina will specialize in women's mental health conditions**, starting with menopause. Menopause is a challenging condition that spans up to a decade[1] and is associated with profound physical, emotional, and social challenges. In Qatar, 49.2% of menopausal women report musculoskeletal pain, 37.2% experience night sweats, 35.4% describe heightened nervousness, and 28.3% face reduced social engagement [2]. Such conditions impair quality of life, increase workplace absenteeism, and diminish productivity, exacerbating economic burdens through healthcare costs and labor market attrition [3], [4].

Impact: Sakina will: (1) **enhance depth in therapy:** psychologists and psychiatrists often lack sufficient time to conduct in-depth therapeutic interventions, especially approaches like psychoanalysis, which require extended sessions. Sakina supports deeper analysis by offloading tasks such as documentation, patient background completion, and symptom tracking between sessions, (2) **ensure continued care and follow-ups:** due to limited availability of mental health providers, patients experience long gaps between sessions, leading to loss of progress, disengagement, and symptom worsening. Sakina fills this gap by keeping patients engaged with personalized therapeutic exercises and follow-ups tailored to their needs and, with patient consent, keeping the doctor in the loop, (3) **automate documentation and assessment:** therapists spend **a significant portion of their time** on documentation, assessment tracking, and patient history updates, reducing the time available for direct patient care. Sakina automates documentation, generating structured progress reports to help clinicians focus on treatment rather than paperwork, (4) **personalize to cultural and gender specific mental health needs:** many therapy approaches use standardized techniques rather than adapting interventions based on individual patient preferences, symptoms, and cultural factors. Sakina focuses on menopausal women's mental health and integrates cultural sensitivity (**Islamic psychology**) and biological factors (hormonal changes, menopause symptom tracking) into its therapeutic approach. (5) Sakina will also be the **first chatbot supporting Islamic Psychology**. Sakina will also help (6) **collect actionable datasets for research**: By logging anonymized symptom progression, treatment responses, and cultural-context data, Sakina addresses gaps in longitudinal menopause studies in Qatar, enabling tailored interventions and policy insights.

How It Works: After each session, Sakina will process an audio (or visual) transcription and generate a structured plan based on the patient's availability until the next appointment. These offline tasks will include: (1) **medical education** tailored to the patient's condition, (2) further **symptom assessment** to validate the doctor's diagnosis, (3) **patient profile completion** (e.g., psychiatric history) and (4) **personalized therapy** aligned with the treatment plan. For personalization, we will build a patient progression model and construct a Graph RAG to keep track of the patient's information. Sakina will then leverage a **multi-agent system** to carry these tasks based on the generated schedule and will synthesize a report to the doctor before the next session detailing progress on the tasks as well as interesting findings (eg., misconceptions and distorted beliefs).

1. INTRODUCTION

1.1 BACKGROUND AND PRELIMINARY DATA: To evaluate stakeholder perspectives on AI integration in mental health care, we conducted four surveys with 129 participants, including mental health providers and patients. The results indicate a high demand for AI-augmented support: 52.94% of providers expressed willingness to use AI chatbots for psychoeducation, patient documentation, and structured symptom tracking. 43.2% of respondents emphasized the necessity of cultural and religious alignment in AI models, while 64.7% supported integrating Islamic psychology principles into AI-driven mental health tools. Patients think AI particularly to be useful for treatment plan personalization based on symptoms and triggers, with over 35% expressing interest in AI-assisted self-help exercises. However, privacy and data security were the most cited concerns, with several respondents reluctant to share sensitive personal information with AI systems. Additionally, skepticism arose regarding AI's ability to interpret emotions and contextual nuances, with concerns that AI might lack the human empathy necessary for effective therapy. Ethical considerations also emerged, with participants advocating for provider oversight (ensuring AI remains an assistive tool), transparent ethical guidelines, and explainable AI mechanisms. These findings reinforce the need for Sakina, an AI-powered chatbot designed to support, not replace, mental health professionals through human-in-the-loop AI, privacy safeguards, culturally adapted responses, and multimodal symptom tracking. We will leverage publicly available datasets (details in Appendix 1). We review Mental Health Support Apps in Appendix 2. **Team Members:** **Dr. Safa Messaoud** is an AI scientist at QCRI, specializing in generative models and reinforcement learning. She is a member of the Fanar team, Qatar's Large Language Model initiative, where she works in the Safety and Alignment team, on developing Arabic-centric and culturally aligned LLMs. This work is supported by an HBKU Signature Research Program (HBKU-OVPR-SRG-02-1), where she serves as Principal Investigator. Her research also explores generalization-robustness tradeoffs in LLM safety, with publications in ECML PKDD 2022 (Oral) and ACL 2023. Safa's prior work on reinforcement learning algorithms has also been featured in top-tier conferences, including CVPR 2020 (Oral), ACM SIGIR 2021, and ICLR 2024. Beyond AI safety and reinforcement learning, Safa has contributed to health intelligence, holding two U.S. patents on problem list generation from electronic medical records and a customized processor for genomic variant calling pipelines. Safa earned her PhD in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign (US), a Master of Science in Computer Engineering from Virginia Tech (US), and both a Master's and Bachelor's in Electrical Engineering and Information Technology from the Technical University of Munich (Germany). She was a visiting scholar at UC Berkeley for a year and completed research internships at IBM Research (Tokyo and New York) during her PhD.

Dr. Leena Babiker is a psychiatry resident and an active researcher in digital mental health, with a strong focus on AI-driven advancements in mental health care. She has an ongoing collaboration with the **Qatar Computing Research Institute (QCRI)**, where she contributes to research on AI applications in psychiatry. She won the **QCRI Generative AI Hackathon**, securing support to develop innovative digital psychiatry tools and successfully secured an **incubation at Qatar Science & Technology Park (QSTP)** to further explore AI-driven solutions in mental health research. Her research on **Telemental Health Care (TMHC) with Hamad Bin Khalifa University (HBKU)** has explored how technology can reduce barriers and stigma in mental health access in Qatar. With numerous peer-reviewed publications in **high-impact journals**, including *Wiley Online Library*, *BMC Genomic Data (Springer Nature)*, *Taylor & Francis Online*, and *Elsevier's The Arts in Psychotherapy*, she bridges the gap between academic research and practical implementation. Dr. Babiker has also contributed to clinical protocol development and interdisciplinary projects focused on digital psychiatry. Additionally, she has led the **Mental Health ECHO Project** in Sudan, working on expanding mental health access through digital platforms and capacity-building initiatives.

Dr. Ines Arous [Assistant Professor at York University, Canada]: Her research is dedicated to advancing the field of human-centered natural language processing (NLP). As a Postdoctoral Researcher at the Mila-Quebec AI Institute and McGill University, she led projects on optimizing the usage of human feedback in language models to improve their explainability and trustworthiness. Her work focuses on designing frameworks that leverage human intelligence to enhance NLP models and tailor them for applications with significant societal impact. Ines got her Ph.D. from the University of Fribourg (Switzerland), specializing in human-AI collaborative approaches for data curation. She developed hybrid human-AI approaches for data collection, evaluation, and model explanation. Her work has been accepted and presented at top conferences, including the **Web Conference, AAAI,** and **ACL**. During her doctoral studies, she completed an internship at Alexa Shopping, Amazon Research. In recognition of her outstanding research, Ines received the prestigious Dimitris N. Chorafas Award for excellence in scientific research in November 2023.

Dr. Hooman Keshavarzi is a licensed clinical psychologist in the state of Illinois, he holds a Doctorate and Masters in Clinical Psychology and a Bachelors of Science – specialist psychology track/minor in Islamic Studies. He currently serves as the program director for the Masters in Counseling Islamic Psychology Program in Doha, Qatar, is a visiting scholar for Ibn Haldun University (Istanbul, Turkey) and adjunct faculty at the Hartford Seminary. He is the founding director of Khalil Center – the first Islamically oriented professional community mental wellness center and largest provider of Muslim mental healthcare in North America. He is also a senior fellow at the International Association for Islamic Psychology (IAIP), conducting research on topics related to Islam, Muslims and Mental Health. Hooman Keshavarzi is an international public speaker and trainer providing education on the intersection of

Islamic studies and behavioral health. Dr. Hooman Keshavarzi has also authored several published academic papers in recognized peer-reviewed journals on integrating Islamic spirituality into modern psychological practice.

Dr. Zaineb Imam is an adult and perinatal psychiatrist with over 30 years of experience and serves as the Division Chief of Women's Mental Health. She specializes in perinatal mental health, including pre- and postnatal mood disorders, psychosis, preconception planning, and complex medication management. Additionally, she consults on general adult mental health conditions such as anxiety, depression, OCD, PTSD, psychosis, and bipolar disorder.Dr. Imam has held academic roles, including Honorary University Fellow at Plymouth University Peninsula Schools of Medicine and Dentistry, where she taught and examined medical students. She also led the General Adult Module for the MRCPsych Course and trained core psychiatry trainees. She has contributed to public education through the Royal College of Psychiatrists' Public Education Engagement Board and was a partner at Psychiatry-UK LLP, a telepsychiatry and ADHD service provider.As Clerkship Director for Medical Students in Women's Mental Health, Dr. Imam leads World Maternal Mental Health Day celebrations and is a member of the International Task Force for the initiative. She also serves as a co-opted executive member of the Perinatal Faculty at the Royal College of Psychiatrists, where she leads international collaborations in perinatal mental health.

1.2 RESEARCH HYPOTHESIS AND OBJECTIVES: **Hypothesis:** (H1) AI for diagnosis and therapy: diagnosis and therapy shouldn't be carried on by an AI assistant independently from the doctor due to: (1) *Task complexity and contextual understanding:* diagnosis and therapy require contextual understanding, clinical intuition, and the ability to integrate complex patient histories, social factors, and comorbidities and AI agents are still not mature in this area. (2) *Uncertainty and nuances in the medical data:* symptoms can be ambiguous, overlapping across multiple conditions. A doctor's experience allows them to interpret unclear cases, consider rare diseases, and adjust their approach based on subtle signs that machines might miss. (3) *Ethical and legal responsibility:* doctors are accountable for medical decisions and patient outcomes. (4) *Patient communication:* A critical part of therapy and treatment involves emotional support, reassurance, and the ability to address patient fears, something that automated systems cannot fully replicate. (5) *Bias and fairness:* biases in training data, lack of generalization across populations, and unpredictable edge cases make full automation risky. Hence, we propose a framework for mental health providers' augmentation. We hypothesise that Sakina will offload tasks like documentation, patient profile completion (e.g., medical, family, educational history), and medical education. It will also augment diagnosis by looking for further evidence in the offline conversation with the patient and augment therapy by conducting offline sessions around topics or a plan that the doctor specifies. (H2) An AI-assistant for mental health providers should be (1) *personalized*: the offline care schedule should take into account the current patient condition, progression, mental state and preferences. A successful execution of this schedule would additionally require the agent to be endowed with a long term memory to remember information across sessions, (2) *proactive*: the agent should steer conversations towards a desired outcome, e.g. resolving symptoms, keeping user engagement (this is in contrast with a reactive approach). To this end, it should be capable of alternating between different tasks for the schedule execution, (3) *multimodal*: to improve accuracy, it should process multi-modal data including speech, text, video and wearables. (4) culturally aware: based on preferences, the schedule can have material relevant to the patient faith. It should also be fluent in the patient dialect. (H3) User segment: Women going through menopause need a close follow-up on their condition for a long period of time (several years). Menopause is a complex phase associated with an increased risk of depression, anxiety, and cognitive changes due to hormonal fluctuations, sleep disturbances, and psychosocial factors [35]. This category will correspond to women born in the 80s, who are likely familiar with technology.

Objectives: (A) Culturally Aware Mental Health Support: the first goal of this project is to develop a library for mental health support inspired by islamic psychology as well as a cultural mental health knowledge graph. (B) Offline care schedule generation: Based on the treatment history and the patient profile, we first propose tasks related to medical education, patient background completion and augmented diagnosis and therapy. The pipelines for these tasks, will leverage medical **knowledge for diagnosis and therapy**, a memory system (**G-RAG**) to store patient relevant information, and a progression model to track the patient's state and progress in the therapy. Then, we will organise these tasks into a schedule taking constraints related to daily availability, emergency of the tasks, and doctor's recommendation and the desired level of variety. For the medical education tasks, we retrieve material from the above mentioned library. (C) Offline care schedule execution: We will have a multi-agent system in the loop with verifiers. We will design agents for different tasks: medical education, patient profile completion, augmented diagnosis and therapy and report generation. We will additionally have different verifiers in the loop checking the reply against different metrics and providing feedback in case of violations. (D) Evaluation: We will evaluate Sakina's effectiveness through automated metrics, clinical validation, comparative benchmarking, and user studies. Automated assessments will measure response quality, coherence, bias, empathy, factual accuracy, and safety. Clinical trials will assess symptom improvement, trust, and adherence, while experts review therapy quality. Sakina will be benchmarked against leading mental health chatbots (e.g., Woebot, Wysa) based on symptom reduction, engagement, empathy, and cultural alignment. User studies will track retention, satisfaction, and long-term improvements, ensuring scientific soundness.

1.3 EXPECTED OUTCOME: Sakina's AI-powered mental health platform directly aligns with Qatar's QRDI priorities in Health, Digital Technology, and Society by addressing critical gaps in menopause-related care, advancing AI-driven healthcare innovation, and supporting Qatar's aging population and societal well-being. Below, we detail how the project's outcomes align with national needs and priorities. (A) Health priority: Menopause-related mental health challenges (anxiety, depression, cognitive decline) are under-addressed in Qatar's healthcare system despite rising cases among women. Sakina bridges this gap with scientifically validated, culturally adapted interventions like CBT and mindfulness, empowering women to seek confidential support and overcoming barriers to mental healthcare. Its scalable model can expand to other underserved groups, strengthening Qatar's mental health infrastructure. (B) Digital technology priority: Qatar seeks to lead in AI-driven healthcare, yet Arabic-language mental health AI models remain scarce. Sakina leverages machine learning and NLP to provide personalized, culturally adapted care, incorporating Arabic dialect support and gender-sensitive design. Its AI continuously refines recommendations based on user interactions, ensuring evidence-based and contextually relevant interventions. This menopause-focused model lays the groundwork for a broader mental health ecosystem, positioning Qatar as a leader in ethical, human-centric AI healthcare innovation. (C) Society priority: With Qatar's growing midlife and senior population [37] , Sakina provides proactive mental health support for menopause-related challenges, enhancing women's workforce productivity and reducing aging stigma. It aligns with Qatar's shift from "elderly" to "senior," promoting dignity, inclusivity, and long-term well-being through culturally adapted interventions. (D) Strategic focus: Sakina strengthens Qatar's leadership in scientific research by fostering digital entrepreneurship, AI talent development, and healthcare partnerships. Globally, it positions Qatar as a pioneer in women's health innovation, aligning with the National Health Strategy (2018–2022) and Qatar National Vision 2030.

2. IMPORTANCE TO QATAR

2.1 ALIGNMENT TO CALL'S RESEARCH PRIORITY AREAS AND RELEVANCE TO THE LOCAL CONTEXT: Sakina's AI-driven mental health platform is rigorously aligned with the research call, addressing all five priority areas through innovative, scalable solutions anchored in Qatar's healthcare needs. By focusing on menopause-related mental health—a high-impact use case where 35.4% of Qatari women report heightened nervousness and 28.3% experience reduced social engagement [2]—Sakina generates critical epidemiological insights (Priority 1), enabling evidence-based policies to mitigate workforce attrition and healthcare costs. Its AI framework, integrating Arabic language models (Fanar LLM) and Islamic psychology, delivers culturally adapted interventions (Priority 2), bridging gaps in locally tailored mental health tools. Leveraging emerging technologies (Priority 3), Sakina combines multimodal data (speech, wearables, and behavioral analytics) to detect early symptom progression and personalize care. The platform's non-invasive, AI-augmented therapies complement clinical practices (Priority 4), offering scalable support for behavioral and neurological challenges. Crucially, Sakina's modular design ensures adaptability beyond menopause: it serves as a blueprint for addressing Priority 5's focus on neurodevelopmental and early-life interventions, such as supporting adolescents with anxiety, by tailoring care to developmental stages, cultural contexts, and biomarker-driven needs. This approach directly supports Qatar's demographic transition, fostering proactive mental health resilience across life stages while advancing QRDI's vision for data-driven, equitable care. By prioritizing culturally grounded innovation and scalable AI, Sakina positions Qatar as a global leader in ethical digital health solutions aligned with the National Health Strategy and Qatar National Vision 2030.

2.2 SOCIAL, HEALTH, ECONOMIC, AND ENVIRONMENTAL IMPACT: (1) Capacity building: By fostering collaborations with QCRI and QSTP, this project enhances local expertise in AI-driven mental health solutions, supporting the development of Qatar-based talent in AI, psychology, and digital healthcare innovation. (2) Private enterprise: Private Enterprise & Innovation: Sakina will serve as the foundation for a scalable digital health startup, promoting entrepreneurship and innovation in Qatar's AI-driven healthcare sector, in line with national health and economic diversification strategies. (3) Economic impact: By addressing mental health gaps and reducing productivity losses from untreated conditions, Sakina is projected to contribute to significant economic savings—estimated at $3.5 billion annually across the GCC—by improving workforce well-being and reducing absenteeism [38]. Furthermore, the broader Middle East mental health market is experiencing substantial growth, projected to exhibit a CAGR of 3.68% from 2024 to 2032, [39] and the mental health apps market in the Middle East & Africa is seeing substantial growth.It is projected to reach a revenue of US$ 728.8 million by 2030, with a CAGR of 14.3% from 2025 to 2030[40]. (4) Health & societal Impact: Providing continuous, AI-augmented mental health support, particularly for underserved groups such as women experiencing menopause-related challenges, will increase mental health accessibility, reduce stigma, and enhance long-term well-being. (5) Cultural impact: By integrating culturally and linguistically tailored AI interventions, Sakina ensures mental health support aligns with local societal norms, making therapy engaging, and effective within the Qatari and broader Arabic-speaking communities.

2.3 COLLABORATION AND INVOLVEMENT OF END-USERS: End-users are central to Sakina's development, ensuring its practicality and effectiveness. Menopausal women, healthcare providers, and mental health professionals actively contribute to the design, testing, and refinement of the platform through focus groups, pilot studies, and feedback loops to address real-world challenges. Expected Benefits: (1) For Women: Access to evidence-based, culturally sensitive mental health support tailored to menopause. (2) For Healthcare Providers: AI-driven tools to enhance patient engagement and treatment adherence. (3) For Policymakers – Data-driven insights on menopause-related mental health trends to inform health policy decisions.

2.4 UTILISATION OF THE JOINT-FUNDER RESOURCES (Not APPLICABLE)

3. WORK PLAN

3.1 METHODOLOGY: This project is organized through four work packages. We aim to leverage recent techniques in AI such as large language models to assist psychotherapists. In WP1, we build clinically and culturally relevant resources to verify and rectify generated content by LLMs. In WP2, we leverage the ressources built in WP1 to design the backbone of a personalised care plan to be executed by Sakina offline in between sessions. In WP3, we propose a multiagent systems to augment diagnosis and care with medical education, patient profile completion and report generation agents. Finally, in WP4, using automatic metrics and human evaluation, we test our assisting tools in simulated and real-world scenarios. All WPs complement each other and address the ultimate goal of providing personalized and culture-aware assistance to patients. A comprehensive overview of all packages interacting with each other is presented in Fig 4 in the Appendix.

3.2 TECHNICAL DESCRIPTION BY WORK PACKAGE

3.2.1. Work Package (WP 1) #: Culturally Aware Library/Knowledge graph for Women Mental Health

Work Package Title	Start Month	End Month
Offline Care Plan Execution	July 2025	June 2026

Participant Name	Role	Efforts Inside Qatar	Efforts Outside Qatar	Total Efforts in Days
Dr. Leena Babiker	PI	2 days/week*42weeks	0	84 days
Dr. Ines Arous	PI	0	2 days/week*18weeks	36 days
Dr. Zainab Imam	PI	1 days/week*20weeks	0	20 days
Dr. Hooman Keshavarzi	Consultant	0.25 day/week*42weeks	0	11 days
RA1	RA	5 days/week*42week	0	210 days
RA2, RA3	RA	5 days/week*4weeks	0	20 days each

Tasks and Deliverables of the WP:

This WP aims to address the lack of cultural alignment in LLMs [41], [42] in mental health applications. We propose to build medical and cultural data sources and structure them as a knowledge graph in tasks 1.1 and 1.2. A knowledge graph is a compelling data structure for these tasks as it allows a structured representation of information where entities (e.g., concepts) are connected through relationships that capture meaningful associations. Moreover, knowledge graphs enable **context-aware augmentation** where LLMs retrieve relevant information before generating responses (task 1.3).

Task 1.1: Building a medical knowledge graph for the mental health of women in menopause [**Duration:** 13 weeks]. **Deliverable:** Medical knowledge graph for diagnosis and therapy for women in menopause. The knowledge graph will encode diagnostic criteria, therapy methodologies, symptom progressions, and treatment pathways. **Methodology:** In this task, we will leverage the most up-to-date and comprehensive existing medical knowledge graphs [43], [44] such as SemMed , PrimeKG [45]and PharmKG [46] to extract the knowledge graph that is relevant to the mental health domain for women in menopause **(Phase 1, 2 weeks)**. We will define the node types in the knowledge graph based on relevant concepts such as the symptoms, the treatment type, and the risk and proactive factors and define the relationships between them (e.g., treatment X helps with symptom Y). **(Phase 2, 3 weeks)** *populate the knowledge graph with mental health-related data:* existing resources from **Phase 1** are designed for the medical domain in general. In this **Phase 2**, we will curate structured mental-health data from clinical guidelines, research papers, and expert input to populate the knowledge graph [47], [48] and encode therapy methodologies (CBT, psychoanalysis, mindfulness) into the graph. Once we have developed a knowledge graph dedicated to mental health with a particular focus on women in menopause, we will integrate our developed knowledge graph with existing methods **(Phase 3, 8 weeks)**. In particular, we will focus on methods that infer diagnosis and treatment pathways based on symptoms and patient history, such as AdaCare [49], GraphCare[50], and BD-Risk [51] and investigate ways to optimize their performance based on our developed knowledge graph. We will evaluate their performance in terms of symptom assessments and therapy plans against those of mental health professionals.

Task 1.2: Cultural knowledge graph (CKG) plug-in [**Duration:** 13 weeks]. Deliverable: The CKG will introduce new entity types to complement the core medical mental health knowledge graph for diagnosis and therapy. These entities will capture cultural attitudes, beliefs, and contextual factors influencing mental health disorders, symptoms, and treatment approaches, with a particular focus on women in menopause. Methodology: Informed by anthropological definitions of culture [52], [53], we aim to build a knowledge graph that goes beyond surface-level cultural representations and includes **context-rich, multi-faceted cultural factors**. Specifically, we will incorporate additional nodes and relationships that capture cultural beliefs, societal norms, community and family influence, and linguistic and expressive variations **(Phase 1, 2 weeks**). We will then populate these defined nodes using research papers, clinical observations, and expert interviews. We will draw inspiration from existing methods for automatically constructing culture-aware knowledge graphs [54], [55] that use clustering techniques to organize concepts into several cultural facets (e.g., traditions, rituals, behaviors). We will leverage existing surveys [56], [57] on mental health in the Arab world to extract relevant concepts. For instance, [56] highlight some commonalities in the Arab world, such as interpreting mental health difficulties through religious and cultural lenses where they are seen as a test from Allah or as a result of Jinn possession or the evil eye. These beliefs are reflected in treatment preferences, where religious and spiritual practices are commonly sought alongside conventional approaches such as cognitive-behavioral therapy (CBT). We will adopt a comprehensive approach where we will collaborate with the Islamic Psychology Department at HBKU alongside psychologists and psychiatrists to refine our knowledge graph. Finally, we will evaluate LLMs with augmented knowledge retrieved from our constructed knowledge graph plugin (e.g., KARD and KGR [58], [59]) using different metrics for appropriateness and bias. We will use recently proposed metrics to evaluate cultural appropriateness, such as surface-level [60], LLM-as-judge [61], and human evaluations. We will also use standard metrics for bias, such as psycholinguistic norms, gender polarity [62] and toxicity [63].
Task 1.3: A library for self-help material for menopause [**Duration:** 16 weeks]. Goal/Deliverable: The library will be used by Sakina to generate personalized in-between session care schedules while accommodating different knowledge levels and including Islamic psychology insights where relevant. The library will be organized by symptoms: mood-related issues (e.g., depression, anxiety, irritability), cognitive symptoms (e.g., brain fog, memory loss, lack of focus), sleep disturbances (e.g., Insomnia), physical symptoms (hot flashes, fatigue, pain), social and emotional well-being (self-esteem, social isolation, relationship changes). The content will be in the form of audios, videos as well psychometric tests covering both educational material (medical, nutrition) and exercises (eg., breathing, physical exercises). We will provide material adapted to different knowledge levels and cultural sensitivities. The library will be provided for self-browsing and for constructing personalized between sessions schedules. Methodology: **(Phase 1, 2 weeks)** Symptom-based content structuring: We will define core symptom categories, create therapy journeys for each symptom progression, and define knowledge levels (beginner, intermediate, advanced). **(Phase 2, 4 weeks)** Content generation and curation: Generate text-based, video, and audio material using GenAI. Integrate validated medical and psychological educational content. Collaborate with Islamic psychology experts for culturally adapted self-help guidance. Collect psychometric tests for symptom tracking & self-assessment. **(Phase 3, 8 weeks)** Adaptive Retrieval System Implementation: Implement AI-driven symptom-based recommendation system, train Retrieval-Augmented Generation (RAG) models to fetch relevant self-help material, and develop personalized content ranking algorithms based on user progress & engagement. **(Phase 4, 2 weeks)** Testing: conduct A/B testing and evaluate retrieval performance.

3.2.2 Work Package (WP 2) #: *Offline Augmented Care Plan Generation*

Work Package Title	Start Month	End Month
Offline Augmented Care Plan Generation	July 2025	July 2026

Participant Name	Role	Efforts Inside Qatar	Efforts Outside Qatar	Total Efforts in Days
Dr. Safa Messaoud	LPI	3 days/weeks*36weeks	0	108 days
Dr. Ines Arous	PI	0	0.5 days/week*12 weeks	6 days
RA 2, RA 3	RA	5 days/week*36 weeks	0	180 days each

Tasks and Deliverables of the WP
The goal of this WP is to generate a personalised care plan to be executed by Sakina offline in between sessions. To this end, we construct a long term memory for storing the patient information (Graph RAG, Task 2.1) and develop a patient progression model (Task 2.2) to quantify the patient cognitive/emotional state and progress in the therapy. We leverage this information to generate a list of customized tasks spanning four categories (educational material, patient profile completion and augmented diagnosis and therapy) and organize these tasks into a schedule (Task 2.3).

Task 2.1: Graph RAG for patient profiles [**Duration:** 8 weeks]. Deliverable: The graph will represent patients' mental health journeys with a comprehensive, evolving structure. It will be used for personalizing the interactions with the ChatBot. Methodology: (**Phase 1, 2 weeks**) Graph structure: Nodes will include current diagnosis, symptoms, triggers and blockers, strength and coping mechanisms, treatment plan (therapy techniques, medications, or lifestyle interventions), progress and milestones and engagement and adherence. The structure will be defined with the help of Dr. Leena (psychiatrist) and Dr. Hooman (psychologist) (**Phase 2, 1 week**) Extracting relevant information from session transcription: We will compare classical NLP pipelines to LLMs. (**Phase 3, 1 week**) Mapping extracted insights to graph RAG nodes. (**Phase 4, 3 weeks**) Validating Graph-RAG for personalized therapy. By interviewing clinicians, we will evaluate the following aspects: (1) how well the knowledge graph's structure aligns with clinical standards. (2) information extraction accuracy, i.e. evaluate how well the NLP pipeline extracts relevant entities (symptoms, diagnosis, treatments, progress) from therapy session transcripts. (3) Speed and accuracy of querying of patient data for therapy decision-making. (4) personalization and adaptability in therapy by measuring patient satisfaction and engagement.

Task 2.2: Patient Progression and Multimodal Fusion **[Duration:** 12 weeks]. Deliverable: The patient progression model will integrate multimodal data sources to track mental health evolution over time. It will track (1) the intensity and frequency of symptoms over time (focus on tasks related to the most prevailing symptom), (2) treatment response and effectiveness (adjust based on the method, e.g., CBT that works best for the patient), (3) Patient adherence and engagement patterns (If adherence drops, Sakina switches to lighter, motivational tasks to rebuild consistency. If engagement remains high, the model gradually introduces more advanced or deeper therapy work.) (4) progress milestones and recovery stages (crisis phase requiring intense support, stabilization phase needing structured self-care, maintenance phase focusing on relapse prevention). In the crisis phase, Sakina schedules frequent check-ins and low-cognitive-load activities. In the maintenance phase, Sakina reduces frequency and shifts toward preventative techniques. Methodology: **(Phase 1, 4 weeks)** Patient progression tracking and therapy adaptation: To develop a robust patient progression model, we will extract insights from patient-reported outcomes, therapy session transcripts, wearable data, and behavioral signals. Symptom trends will be monitored through self-reported check-ins, therapist notes, and physiological signals such as heart rate variability from wearables. Bayesian time-series models will estimate symptom intensity fluctuations and identify dominant issues requiring priority intervention. Treatment response will be assessed by analyzing task feedback, session summaries, and sentiment shifts, allowing therapy plans to be dynamically adjusted. Behavioral patterns, including task completion rates and engagement levels, will guide therapy intensity, ensuring that low motivation triggers lighter interventions while sustained engagement allows for deeper therapy work. **(Phase 2, 8 weeks)** Multimodal Data Fusion for Personalized Insights: advanced data fusion techniques will be explored to integrate text (session transcripts), audio (voice tone and speech patterns), physiological data (heart rate variability, sleep patterns, stress levels), and behavioral engagement metrics into a holistic patient state representation. Graph neural networks (GNNs) will be applied for structured multimodal reasoning, while attention-based transformers will facilitate hierarchical feature integration. Variational autoencoders (VAEs) will support latent space alignment across modalities. Additionally, Tensor Fusion Networks (TFN) and Cross-modal Transformers will be explored to capture interdependencies between signals. State-space models will track phase transitions, ensuring therapy evolves dynamically based on patient stress levels or crisis phases. This approach will optimize therapy personalization, enabling Sakina to anticipate needs, track recovery, and recommend adaptive interventions tailored to individual mental health journeys.

Task 2.3: Offline care plan generation algorithm [**Duration:** 16 weeks]. Deliverable: The offline care plan is a schedule to be executed by the ChatBot in between sessions. It will cover augmented diagnosis and therapy tasks as well as medical education and patient profile completion. Methodology: **(Phase 1, 6 weeks)** Generate a list of tasks. For augmented diagnosis: query the medical knowledge graph (Task 1.1) to identify undiagnosed but related symptoms. For augmented therapy: query the knowledge graph to get the sequence of steps of the therapy approach recommended by the doctor and query the Graph RAG (Task 2.1) to extract the relevant blockers and strengths to the therapy goal. For medical education: identify the content based on symptom severity and progress. For patient medical history: identify the missing patient background items based on the Graph RAG content. **(Phase 2, 1 week)** Retrieve appropriate material and exercises from the library based on primary symptoms, patient knowledge level, and cultural fitness. (**Phase 3, 8 weeks**) Convert the list of tasks into a schedule given constraints on the daily availability until the next appointment, emotional readiness, engagement, therapeutic sequence and progression, doctor's recommendations, and priorities. Learn the optimal schedule using constrained reinforcement learning [64]. To this end, we will explore generating offline recommendation trajectories from the sessions by leveraging the Alexander Street datasets [65] or via synthetic data generation through patient simulations following PATIENT-Ψ[66] framework.

3.2.3 Work Package (WP 3) #: *Offline Care Plan Execution*

Work Package Title	Start Month	End Month
Offline Care Plan Execution	July 2026	May 2027

Participant Name	Role	Efforts Inside Qatar	Efforts Outside Qatar	Total Efforts in Days
Dr. Safa Messaoud	LPI	3 days/week*48 weeks	0	144
RA 1	RA	5 days/week*48 weeks	0	240
RA 2, R3	RA	4 days/week*48 weeks	0	192 each

Tasks and Deliverables of the WP

Sakina's offline care plan will be managed by a multi-agent system, including agents for medical education, patient profile completion, diagnosis, therapy, scheduling, engagement, and report generation. This setup ensures specialization, scalability, and efficiency, allowing agents to handle tasks in parallel while being fine-tuned for specific functions. Agents will be trained using supervised and reinforcement learning (RL), with coordination explored through hierarchical RL and multi-agent RL (MARL) [67] to optimize user engagement, task completion, and symptom resolution. The system will track patient condition, engagement, and task progress, with an action space tailored to general and agent-specific actions (e.g., therapy strategies like CBT vs Psychoanalysis). A patient simulator will be used for online RL training, making the system adaptable and effective in delivering AI-driven mental health support.

Task 3.0: **Patient simulator** (**Duration:** 10 weeks). Deliverables: The patient simulator will serve as the core environment for training and evaluating Sakina's multi-agent reinforcement learning system. It will generate realistic patient interactions with diverse personalities, adherence levels, and medical conditions, allowing agents to learn how to handle real-world variability before deployment. Methodology: We will first generate different patient profiles and treatment scenarios (eg., islamic psychology/western). Then, we will collect data by deploying the simulated patients in our system. Task 3.1: **Medical education agent** (**Duration:** 2 weeks). Deliverable: The medical education agent (MEA) will provide personalized, structured educational content based on the patient's diagnosis and symptoms, severity of the condition and progression, patient's knowledge level (Beginner, Intermediate, Advanced) and cultural sensitivity (e.g., faith-based perspectives, language, and societal norms). Methodology: we will implement the following capabilities: (1) retrieval of additional material from the library (task 1.3) or medical knowledge graph (task 1.1) to answer user questions, (2) access to the user G-RAG (task 2.1) to retrieve related personal experiences to the task for a more personalized context, (3) communication with the engagement agent to log knowledge progression, degree of engagement and knowledge holes, (4) communication with the patient profile completion agent to transfer background information to be used to update the G-RAG, and with the report generation agent to provide a summary of the task execution. Task 3.2: **Patient profile completion agent** (**Duration:** 2 weeks) Deliverable: The patient profile completion agent role is to gather background information about the patient (family, education, relationships, medical). The agent is primarly responsible for tasks requested by the mental health provider. It can however also initiate additional material collection when in need. Methodology: We will implement: (1) communication with the G-RAG to retrieve relevant information to the background task to executed, (2) extraction of causality relationships to store/update the collected background material in a structured way in the G-RAG (eg., blocker and blocker root causes), (3) communication with the engagement agent to assess the user cognitive/mental state, (4) communication with the scheduler agent to add additional tasks on background material and with the report generation agent to provide a summary of the task execution. Task 3.3: **Augmented diagnosis agent** (**Duration:** 2 weeks). Deliverable: The augmented diagnosis agent role is to find evidence for the doctor diagnosis and assess different symptoms. Methodology: We will implement the following capabilities: (1) the diagnosis agent will receive the doctor diagnosis and current symptoms, it will query the medical knowledge graph for all the symptoms related to that disorder, other disorders with similar symptoms and the cultural knowledge graph for additional information how how these symptoms are expressed, as well as medical history from the G-RAG, (2) the agent will analyse different modalities (peach tone, facial expressions, text sentiment, and wearable data) to detect evidence of different symptoms. To achieve this, we will have a library of classifiers for different symptoms/disorders (e.g., a depression intensity classifier trained on D4 [27]) that can be deployed. For the symptoms/disorders that we don't have a dataset for, we will rely on the agent's judgement and reasoning. (3) The augmented diagnosis agent also can walk the user through psychometric tests or request tests from the scheduler. (4) It will communicate with the report generation agent to provide a summary of the task execution. Task 3.4: **Augmented therapy agent** (**Duration:** 8 weeks). Deliverable: The augmented therapy agent will execute personalized therapy exercises. Methodology: The agent will have the following capabilities: (1) It will extract knowledge about the patient from G-RAG whenever is necessary to leverage it in the response generation. (2) It will extract the guidelines of the therapy approach from the medical knowledge graph. (3) It will communicate with the engagement agent to adapt the conversation based on the patient interaction and with the report generation agent to provide a summary of the task execution. We will pre-train this agent using reinforcement learning on the ESConv dataset[25]. Task 3.5: **Scheduler agent** (**Duration:** 3 weeks). Deliverable: the scheduler agent collects the proposed additional tasks by the other agents and communicates with the scheduler (Task 2.3) to propose an adjusted schedule. Methodology: The scheduler agent will communicate the engagement level to different tasks and the user preferences conveyed by the engagement agent to the scheduler. It will notify the doctor through the report generation agent to endorse the proposed schedule. Task 3.6: **Engagement agent** (**Duration:** 2 weeks) Deliverable: The engagement agent monitors adherence and the patient's cognitive and emotional state. Methodology: (1) The agent will construct the patient's current state by running sentiment analysis [25] on user responses, analyse the wearable data and track the level of focus on the task. The state will be fed into the reinforcement learning algorithm to decide on the next action to take. (2) The agent will task completion and task engagement history to communicate it to the scheduler. (3) It will also infer preferences and progression and update the progression model (Task 2.2) and the G-RAG (Task 2.1) with this information. Task 3.7: **Report generation agent** (**Duration:** 2 weeks) Deliverable: The report generation agent will summarize the offline schedule execution to the doctor and will generate a Json file to visualize different metrics on the doctor dashboard. Methodology: The report summarizes diagnostic findings, therapy progress, engagement trends, and recommendations, ensuring a data-driven, evidence-based approach to patient care. It also computes metrics that will be visualized in the patient/doctor dashboard. Task 3.8: **Verifiers** (**Duration:** 8 weeks). Deliverable: To ensure that Sakina outputs only safe, ethical, and clinically sound utterances, we will implement several layers of verifiers before an utterance is generated. Methodology: The verifiers are LLMs prompt to evaluate content for safety, medical accuracy, ethical considerations, and personalization. (1) **Clinical accuracy verifier:** ensures utterances align with evidence-based medical knowledge. (2) **Ethical and bias verifier:** detects harmful stereotypes, discriminatory language, or unethical advice. (3) **Safety and crisis verifier:** scans for crisis-related indicators (self-harm, suicidal intent, extreme distress). If distress is detected, triggers escalation to the crisis support agent. (4) **Cultural sensitivity verifier:** ensures responses align with the cultural, religious, and linguistic background of the user. **(5) Personalization and tone verifier:** ensures responses are not robotic but warm, empathetic, and user-appropriate. **(6) Consistency and logical coherence verifier:** avoid divergence.

3.2.4 Work Package (WP 4) #: *Evaluation Metrics*

Work Package Title	Start Month	End Month
Evaluation Metrics	May 2027	July 2027

Participant Name	Role	Efforts Inside Qatar	Efforts Outside Qatar	Total Efforts in Days
Dr. Ines Arous	PI	0	2 days/week*8 weeks	16 days
Dr. Leena Babiker	PI	2days/week*8 weeks	0	16 days
Dr. Zainab Imam	PI	2days/week*8 weeks	0	16 days
Dr. Hooman Keshavarzi	Consultant	1 day/week*8 weeks	0	8 days
RA1, R2, R3	RA	5days/week*8 weeks	0	40 days each

Tasks and Deliverables of the WP

We will explore (1) automated evaluation metrics to measure response quality, coherence, and accuracy. (2) Clinical validation via trials to assesses real-world therapeutic effectiveness, (3) Comparative benchmarking via comparing Sakina to other mental health chatbots. (4) We will design user studies to assess user engagement and satisfaction

Task 4.1 **Automated metrics** (**Duration**: 4 weeks) include: (1) Response relevance (BLEU, ROUGE, BERTScore), (2) Coherence and consistency (Dialog Entailment Score), (3) Toxicity and bias detection (Perspective API, Fairness Score), (4) Empathy and conversational engagement (Empathic response score), (5) Factual Accuracy (medical knowledge graph validation) and (5) Safety and crisis detection recall. Task 4.2 **Clinical validation** (**Duration**: 8 weeks): Sakina's will assess its impact on therapy augmentation, patient engagement, and adherence through expert evaluations, therapist-in-the-loop assessments, and real-world observational studies. (1) Expert Evaluation: Psychiatrists and psychologists will review Sakina's therapy recommendations, conversational quality, and adherence to clinical guidelines by analyzing session logs, reports, and therapeutic interactions. (2) Therapist-in-the-Loop Assessment: Mental health professionals will integrate Sakina into their practice to evaluate its role in enhancing patient engagement, reducing documentation workload, and supporting therapeutic goals. (3) Real-World Observational Studies: Patients will interact with Sakina between therapy sessions to assess its effectiveness in providing continuous, structured support. Task 4.3 **Comparative benchmarking** (**Duration**: 4 weeks): Sakina will be benchmarked against leading mental health chatbots (e.g., Woebot, Wysa) to evaluate its effectiveness in symptom reduction, user engagement, retention, response quality, empathy, trustworthiness, and cultural alignment (Arabic dialect, religious considerations). This study will advance Islamic psychology by empirically assessing the impact of faith-based mental health interventions in AI-driven therapy. It will highlight how culturally and religiously sensitive AI improves mental health outcomes for Muslim users, fosters trust in AI-assisted therapy, and enhances user engagement, therapy adherence, and overall well-being. Task 4.4 **User experience** (**Duration**: 4 weeks)**:** In this task, we will leverage existing work in human computation for assessing mental health support systems [68], [69], [70]and evaluate the ease of use of Sakina using the system usability scale questionnaire (SUS) [71]. Additionally, we will asses users' satisfaction with the responses by evaluating Sakina's generated answers compared with other tools in terms of fluency, conversation depth and empathy similar to existing mental health support systems.

3.3 ESTIMATED TIMELINE

WPs & Tasks	Year 1												Year 2											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
WP1	x	x	x	x	x	x	x	x	x	x	x													
T1.1	x	x	x																					
T1.2				x	x	x																		
T1.3	x	x	x	x	x	x	x	x	x	x	x													
WP2	x	x	x	x	x		x	x	x			x												
T2.1	x	x																						
T2.2			x	x	x																			
T2.3							x	x	x			x												
WP3													x	x	x	x	x	x	x	x	x	x	x	x
T3.0													x	x	x									
T3.1-3.3															x	x								
T3.4																	x	x						
T3.5-3.7																		x	x					
T3.8																		x	x					
T3.9																				x	x			
W 4 (parallel execution)																						x	x	x

4. PROJECT MANAGEMENT AND SYNERGY: Dr. Safa and Dr. Leena will oversee the coordination on the computer science and medical parts. They have been working together for the last six months and have participated with Sakina's idea in QU Health hackathon last november and recently Al Fikra national entrepreneurship competition (8 weeks program), where they refined the idea into the current proposal. Dr. Ines visited QCRI twice in the past year and has been actively shaping the proposed workpackage with Dr. Safa over the last months. In the first phase of the project, Dr. Leena will coordinate with Dr. Hooman and Dr. Zaineb on the constructing the ontology of the medical knowledge graph, the cultural plug-in and the structure of the Graph-RAG. Then, Dr. Safa will coordinate with Dr. Ines on completing the knowledge graphs and proceeds with the care-plan schedule generation and execution (WP2,3) and Dr. Leena will move to working together with Dr. Hooman and Dr. Zaineb on the library. In the last phase, the team will work together on the evaluation work package. The project is low risk as we will leverage publicly available datasets and QCRI infrastructure to train the AI models and we will use resources from the islamic psychology department and Sidra Medicine's Menopause Clinic under Dr. Hooman and Dr. Zaineb guidance to construct the library and for evaluating the model. **4.1 Synergy and Relevance of the research team:** The research team brings complementary expertise essential for Sakina's development. Dr. Safa and Dr. Ines are AI specialists—Dr. Safa's expertise in reinforcement learning will guide multi-agent system design and scheduling, while her experience in large language model cultural alignment from the Fanar project will support Sakina's adaptation. Dr. Ines, an expert in human-centered NLP, will contribute to knowledge graph construction and model evaluation. Dr. Leena and Dr. Zaineb, both psychiatrists, and Dr. Hooman, a psychologist, provide medical and cultural expertise—Dr. Zaineb specializes in women's mental health, while Dr. Hooman focuses on Islamic psychology. Dr. Leena will coordinate with Dr. Zaineb on medical knowledge graph ontologies and model validation and with Dr. Hooman on cultural integration and the Islamic psychology library. **4.2 Ethical and regulatory requirementL:** This research will strictly follow Qatar's Ministry of Public Health (MoPH) and IRB guidelines, ensuring participant safety, data confidentiality, and ethical integrity under Good Clinical Practice (ICH-GCP). All ethical approvals and informed consent will be obtained before research begins. Data collection will be non-invasive and observational, with no clinical interventions or drug testing. If integrated into clinical settings, additional MoPH approvals will be secured. Sidra Medicine, a key partner, ensures strict ethical oversight in alignment with women's mental health research standards. **4.3 Dissimilation and exploitation of results:** This grant will establish Sakina, the first culturally aware chatbot integrating Islamic psychology for augmented mental health diagnosis and therapy. The project aims to publish at least three papers in top AI and psychiatry journals and engage with Sidra, HBKU, and potentially Al-Mujadilah to develop its knowledge library. Sidra Medicine's Menopause Clinic will serve as a pilot site, leveraging its expertise in women's mental health. For long-term sustainability, the team will explore commercialization through QSTP and health-tech accelerators, secure intellectual property protections, and pursue strategic partnerships for funding and scaling. Sakina aims to pioneer AI-driven mental health innovation in Qatar and beyond.

3. REFERENCES

[1] N. Santoro, C. Roeca, B. A. Peters, and G. Neal-Perry, "The Menopause Transition: Signs, Symptoms, and Management Options," *J. Clin. Endocrinol. Metab.*, vol. 106, no. 1, pp. 1–15, Jan. 2021, doi: 10.1210/clinem/dgaa764.

[2] A. Bener and A. Falah, "A measurement-specific quality-of-life satisfaction during premenopause, perimenopause and postmenopause in Arabian Qatari women," *J. -Life Health*, vol. 5, no. 3, p. 126, 2014, doi: 10.4103/0976-7800.141190.

[3] W. H. Utian, "Psychosocial and socioeconomic burden of vasomotor symptoms in menopause: A comprehensive review," *Health Qual. Life Outcomes*, vol. 3, no. 1, p. 47, Dec. 2005, doi: 10.1186/1477-7525-3-47.

[4] A. El Hajj *et al.*, "Menopausal symptoms, physical activity level and quality of life of women living in the Mediterranean region," *PLOS ONE*, vol. 15, no. 3, p. e0230515, Mar. 2020, doi: 10.1371/journal.pone.0230515.

[5] "Fanar." Accessed: Mar. 11, 2025. [Online]. Available: https://fanar.qa/en#about

[6] "MindShift® CBT App," Anxiety Canada. Accessed: Mar. 11, 2025. [Online]. Available: https://www.anxietycanada.com/resources/mindshift-cbt/

[7] "How We Feel." Accessed: Mar. 11, 2025. [Online]. Available: https://howwefeel.org/

[8] "Chat with all of your Course Materials | Unstuck AI." Accessed: Mar. 11, 2025. [Online]. Available: https://unstuckstudy.com

[9] "Muraqaba App - Muslim Mindfulness & Wellbeing." Accessed: Mar. 11, 2025. [Online]. Available: https://www.muraqaba.app/

[10] "Muslim Moodfit," Muslim Moodfit. Accessed: Mar. 11, 2025. [Online]. Available: https://muslimmoodfit.com/

[11] "Sabr | App." Accessed: Mar. 11, 2025. [Online]. Available: https://sabrapp.com/

[12] "Woebot Health," Woebot Health. Accessed: Mar. 11, 2025. [Online]. Available: https://woebothealth.com/

[13] "Wysa - Everyday Mental Health," Wysa - Everyday Mental Health. Accessed: Jul. 11, 2024. [Online]. Available: https://www.wysa.com/

[14] "Pi, your personal AI." Accessed: Mar. 11, 2025. [Online]. Available: https://pi.ai/onboarding

[15] "Syrona Health | Gender-Inclusive Employee-Wellbeing Benefits." Accessed: Mar. 11, 2025. [Online]. Available: https://syronahealth.com/

[16] "Sonia AI Emotional Support." Accessed: Mar. 11, 2025. [Online]. Available: https://www.soniahealth.com/

[17] "Online Therapy & Counseling For Women & Mental Health Services by LunaJoy," Mental Health Therapy For Women Covered by Insurance. Accessed: Mar. 11, 2025. [Online]. Available: https://hellolunajoy.com

[18] "Manifest: 24/7 Wellness on the App Store." Accessed: Mar. 11, 2025. [Online]. Available: https://apps.apple.com/ca/app/manifest-24-7-wellness/id6463312362

[19] "balance - balance app." Accessed: Mar. 11, 2025. [Online]. Available: https://www.balance-menopause.com/balance-app/

[20] "Caria: Advancing Women's Post-Reproductive Health," Caria: Menopause Simplified. Accessed: Mar. 11, 2025. [Online]. Available: https://hellocaria.com/

[21] "Menopause support & relief," Stella. Accessed: Mar. 11, 2025. [Online]. Available: https://www.onstella.com/

[22] "EmoCareAI/Psych8k · Datasets at Hugging Face." Accessed: Mar. 11, 2025. [Online]. Available: https://huggingface.co/datasets/EmoCareAI/Psych8k

[23] A. Welivita and P. Pu, "Curating a Large-Scale Motivational Interviewing Dataset Using Peer Support Forums," in *Proceedings of the 29th International Conference on Computational Linguistics*, N. Calzolari, C.-R. Huang, H. Kim, J. Pustejovsky, L. Wanner, K.-S. Choi, P.-M. Ryu, H.-H. Chen, L. Donatelli, H. Ji, S. Kurohashi, P. Paggio, N. Xue, S. Kim, Y. Hahm, Z. He, T. K. Lee, E. Santus, F. Bond, and S.-H. Na, Eds., Gyeongju, Republic of Korea: International Committee on Computational Linguistics, Oct. 2022, pp. 3315–3330. Accessed: Mar. 11, 2025. [Online]. Available: https://aclanthology.org/2022.coling-1.293/

[24] Z. Zheng, L. Liao, Y. Deng, L. Qin, and L. Nie, "Self-chats from Large Language Models Make Small Emotional Support Chatbot Better," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, L.-W. Ku, A. Martins, and V. Srikumar, Eds., Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 11325–11345. doi: 10.18653/v1/2024.acl-long.611.

[25] S. Liu *et al.*, "Towards Emotional Support Dialog Systems," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, C. Zong, F. Xia, W. Li, and R. Navigli, Eds., Online: Association for Computational Linguistics, Aug. 2021, pp. 3469–3483. doi: 10.18653/v1/2021.acl-long.269.

[26] M. Zhang *et al.*, "CBT-Bench: Evaluating Large Language Models on Assisting Cognitive Behavior Therapy,"

Jan. 26, 2025, *arXiv*: arXiv:2410.13218. doi: 10.48550/arXiv.2410.13218.

[27] B. Yao *et al.*, "D4: a Chinese Dialogue Dataset for Depression-Diagnosis-Oriented Chat," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds., Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 2438–2459. doi: 10.18653/v1/2022.emnlp-main.156.

[28] A. Cohan, B. Desmet, A. Yates, L. Soldaini, S. MacAvaney, and N. Goharian, "SMHD: a Large-Scale Resource for Exploring Online Language Usage for Multiple Mental Health Conditions," in *Proceedings of the 27th International Conference on Computational Linguistics*, E. M. Bender, L. Derczynski, and P. Isabelle, Eds., Santa Fe, New Mexico, USA: Association for Computational Linguistics, Aug. 2018, pp. 1485–1497. Accessed: Mar. 11, 2025. [Online]. Available: https://aclanthology.org/C18-1126/

[29] G. Gkotsis *et al.*, "Characterisation of mental health conditions in social media using Informed Deep Learning," *Sci. Rep.*, vol. 7, no. 1, p. 45141, Mar. 2017, doi: 10.1038/srep45141.

[30] C. Yin *et al.*, "MDD-5k: A New Diagnostic Conversation Dataset for Mental Disorders Synthesized via Neuro-Symbolic LLM Agents," Dec. 26, 2024, *arXiv*: arXiv:2408.12142. doi: 10.48550/arXiv.2408.12142.

[31] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "GoEmotions: A Dataset of Fine-Grained Emotions," 2020, *arXiv*. doi: 10.48550/ARXIV.2005.00547.

[32] J. Gratch *et al.*, "The Distress Analysis Interview Corpus of human and computer interviews," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC`14)*, N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, and S. Piperidis, Eds., Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, pp. 3123–3128. Accessed: Mar. 11, 2025. [Online]. Available: https://aclanthology.org/L14-1421/

[33] Y. Chu, L. Liao, Z. Zhou, C.-W. Ngo, and R. Hong, "Towards Multimodal Emotional Support Conversation Systems," Oct. 19, 2024, *arXiv*: arXiv:2408.03650. doi: 10.48550/arXiv.2408.03650.

[34] A. Bagher Zadeh, P. P. Liang, S. Poria, E. Cambria, and L.-P. Morency, "Multimodal Language Analysis in the Wild: CMU-MOSEI Dataset and Interpretable Dynamic Fusion Graph," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Melbourne, Australia: Association for Computational Linguistics, 2018, pp. 2236–2246. doi: 10.18653/v1/P18-1208.

[35] M. M. Murphy, M. A. Verjee, A. Bener, and L. M. Gerber, "The hopeless age? A qualitative exploration of the experience of menopause in Arab women in Qatar," *Climacteric*, vol. 16, no. 5, pp. 550–554, Oct. 2013, doi: 10.3109/13697137.2013.771330.

[36] J. A. Salinas-Perez, S. Ghuloum, M. R. Gutierrez-Colosia, N. Bagheri, and L. Salvador-Carulla, "Analysis of the mental health service provision in Qatar: Progressing in community mental health in a Middle East country," *Front. Psychiatry*, vol. 13, p. 1095788, Dec. 2022, doi: 10.3389/fpsyt.2022.1095788.

[37] UNFPA, "The rights and wellbeing of older persons in Qatar," arabstates. Accessed: Mar. 11, 2025. [Online]. Available: https://arabstates.unfpa.org/sites/default/files/pub-pdf/country_profile_-_qatar_10-1-2024.pdf

[38] PricewaterhouseCoopers, "PwC Middle East report: Why GCC Governments Should Invest More in Mental Health," PwC. Accessed: Mar. 11, 2025. [Online]. Available: https://www.pwc.com/m1/en/media-centre/2022/pwc-middle-east-report-why-gcc-governments-should-invest-in-mental-health.html

[39] "Middle East Mental Health Market Size, Share, Report 2032." Accessed: Mar. 11, 2025. [Online]. Available: https://www.imarcgroup.com/middle-east-mental-health-market

[40] "Mental Health Apps Market Size, Share & Global Report [2030]." Accessed: Mar. 11, 2025. [Online]. Available: https://www.fortunebusinessinsights.com/mental-health-apps-market-109012

[41] T. Naous, C. Hokayem, and H. Hajj, "Empathy-driven Arabic Conversational Chatbot," in *Proceedings of the Fifth Arabic Natural Language Processing Workshop*, I. Zitouni, M. Abdul-Mageed, H. Bouamor, F. Bougares, M. El-Haj, N. Tomeh, and W. Zaghouani, Eds., Barcelona, Spain (Online): Association for Computational Linguistics, Dec. 2020, pp. 58–68. Accessed: Jun. 12, 2024. [Online]. Available: https://aclanthology.org/2020.wanlp-1.6

[42] W. Wang *et al.*, "Not All Countries Celebrate Thanksgiving: On the Cultural Dominance in Large Language Models," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, L.-W. Ku, A. Martins, and V. Srikumar, Eds., Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 6349–6384. doi: 10.18653/v1/2024.acl-long.345.

[43] H. Cui *et al.*, "A Survey on Knowledge Graphs for Healthcare: Resources, Application Progress, and Promise," presented at the ICML 3rd Workshop on Interpretable Machine Learning in Healthcare (IMLH), Jul. 2023. Accessed: Mar. 07, 2025. [Online]. Available: https://openreview.net/forum?id=CZCktJoBRh#all

[44] S. Freidel and E. Schwarz, "Knowledge graphs in psychiatric research: Potential applications and future

perspectives," *Acta Psychiatr. Scand.*, vol. 151, no. 3, pp. 180–191, Mar. 2025, doi: 10.1111/acps.13717.
[45] P. Chandak, K. Huang, and M. Zitnik, "Building a knowledge graph to enable precision medicine," *Sci. Data*, vol. 10, no. 1, p. 67, Feb. 2023, doi: 10.1038/s41597-023-01960-3.
[46] S. Zheng *et al.*, "PharmKG: a dedicated knowledge graph benchmark for bomedical data mining," *Brief. Bioinform.*, vol. 22, no. 4, p. bbaa344, Jul. 2021, doi: 10.1093/bib/bbaa344.
[47] A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Booksmith Publishing LLC, 2021.
[48] "ICD-11." Accessed: Mar. 07, 2025. [Online]. Available: https://icd.who.int/en
[49] L. Ma *et al.*, "AdaCare: Explainable Clinical Health Status Representation Learning via Scale-Adaptive Feature Extraction and Recalibration," *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 01, Art. no. 01, Apr. 2020, doi: 10.1609/aaai.v34i01.5427.
[50] P. Jiang, C. Xiao, A. R. Cross, and J. Sun, "GraphCare: Enhancing Healthcare Predictions with Personalized Knowledge Graphs," presented at the The Twelfth International Conference on Learning Representations, Oct. 2023. Accessed: Mar. 07, 2025. [Online]. Available: https://openreview.net/forum?id=tVTN7Zs0ml
[51] D. Lee *et al.*, "Detecting Bipolar Disorder from Misdiagnosed Major Depressive Disorder with Mood-Aware Multi-Task Learning," in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, K. Duh, H. Gomez, and S. Bethard, Eds., Mexico City, Mexico: Association for Computational Linguistics, Jun. 2024, pp. 4954–4970. doi: 10.18653/v1/2024.naacl-long.278.
[52] P. Bourdieu, *Outline of a Theory of Practice*. Cambridge University Press, 1972.
[53] C. Geertz, *The Interpretation Of Cultures*. Basic Books, 1973.
[54] A. Deshpande, D. Ruiter, M. Mosbach, and D. Klakow, "StereoKG: Data-Driven Knowledge Graph Construction For Cultural Knowledge and Stereotypes," in *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)*, K. Narang, A. Mostafazadeh Davani, L. Mathias, B. Vidgen, and Z. Talat, Eds., Seattle, Washington (Hybrid): Association for Computational Linguistics, Jul. 2022, pp. 67–78. doi: 10.18653/v1/2022.woah-1.7.
[55] T.-P. Nguyen, S. Razniewski, A. Varde, and G. Weikum, "Extracting Cultural Commonsense Knowledge at Scale," in *Proceedings of the ACM Web Conference 2023*, Austin TX USA: ACM, Apr. 2023, pp. 1907–1917. doi: 10.1145/3543507.3583535.
[56] H. alHarbi, P. Farrand, and K. Laidlaw, "Understanding the beliefs and attitudes towards mental health problems held by Muslim communities and acceptability of Cognitive Behavioral Therapy as a treatment: systematic review and thematic synthesis," *Discov. Ment. Health*, vol. 3, no. 1, p. 26, Nov. 2023, doi: 10.1007/s44192-023-00053-2.
[57] F. Fekih-Romdhane *et al.*, "Cross-cultural comparison of mental illness stigma and help-seeking attitudes: a multinational population-based study from 16 Arab countries and 10,036 individuals," *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 58, no. 4, pp. 641–656, Apr. 2023, doi: 10.1007/s00127-022-02403-x.
[58] X. Guan *et al.*, "Mitigating large language model hallucinations via autonomous knowledge graph-based retrofitting," in *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, in AAAI'24/IAAI'24/EAAI'24, vol. 38. AAAI Press, Feb. 2024, pp. 18126–18134. doi: 10.1609/aaai.v38i16.29770.
[59] M. Kang, S. Lee, J. Baek, K. Kawaguchi, and S. J. Hwang, "Knowledge-Augmented Reasoning Distillation for Small Language Models in Knowledge-Intensive Tasks," presented at the Thirty-seventh Conference on Neural Information Processing Systems, Nov. 2023. Accessed: Mar. 05, 2025. [Online]. Available: https://openreview.net/forum?id=xJLEQQrFia&noteId=u1vq1PIP9G
[60] S. Arora, M. Karpinska, H.-T. Chen, I. Bhattacharjee, M. Iyyer, and E. Choi, "CaLMQA: Exploring culturally specific long-form question answering across 23 languages," Jul. 03, 2024, *arXiv*: arXiv:2406.17761. doi: 10.48550/arXiv.2406.17761.
[61] L. Zheng *et al.*, "Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena," Dec. 24, 2023, *arXiv*: arXiv:2306.05685. doi: 10.48550/arXiv.2306.05685.
[62] J. Dhamala *et al.*, "BOLD: Dataset and Metrics for Measuring Biases in Open-Ended Language Generation," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, in FAccT '21. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 862–872. doi: 10.1145/3442188.3445924.
[63] E. Sheng, K.-W. Chang, P. Natarajan, and N. Peng, "The Woman Worked as a Babysitter: On Biases in Language Generation," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language*

Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), K. Inui, J. Jiang, V. Ng, and X. Wan, Eds., Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 3407–3412. doi: 10.18653/v1/D19-1339.

[64] H. Li, Z. Wan, and H. He, "Constrained EV Charging Scheduling Based on Safe Deep Reinforcement Learning," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2427–2439, May 2020, doi: 10.1109/TSG.2019.2955437.

[65] "Behavioral & Mental Health," Alexander Street. Accessed: Mar. 11, 2025. [Online]. Available: https://alexanderstreet.com/discipline/behavioral-mental-health

[66] R. Wang *et al.*, "PATIENT-ψ: Using Large Language Models to Simulate Patients for Training Mental Health Professionals," in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Y. Al-Onaizan, M. Bansal, and Y.-N. Chen, Eds., Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 12772–12797. doi: 10.18653/v1/2024.emnlp-main.711.

[67] D. Huh and P. Mohapatra, "Multi-agent Reinforcement Learning: A Comprehensive Survey," Jul. 03, 2024, *arXiv*: arXiv:2312.10256. doi: 10.48550/arXiv.2312.10256.

[68] M. De Choudhury, S. S. Sharma, T. Logar, W. Eekhout, and R. C. Nielsen, "Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, Portland Oregon USA: ACM, Feb. 2017, pp. 353–369. doi: 10.1145/2998181.2998220.

[69] S. R. Pendse *et al.*, "From Treatment to Healing:Envisioning a Decolonial Digital Mental Health," in *CHI Conference on Human Factors in Computing Systems*, New Orleans LA USA: ACM, Apr. 2022, pp. 1–23. doi: 10.1145/3491102.3501982.

[70] R. Zhang, J. Eschler, and M. Reddy, "Online Support Groups for Depression in China: Culturally Shaped Interactions and Motivations," *Comput. Support. Coop. Work CSCW*, vol. 27, no. 3, pp. 327–354, Dec. 2018, doi: 10.1007/s10606-018-9322-4.

[71] john Brooke, "SUS: A 'Quick and Dirty' Usability Scale," in *Usability Evaluation In Industry*, CRC Press, 1996.

[72] Y. Fei, Y. Hou, Z. Chen, and A. Bosselut, "Mitigating Label Biases for In-context Learning," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, A. Rogers, J. Boyd-Graber, and N. Okazaki, Eds., Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 14014–14031. doi: 10.18653/v1/2023.acl-long.783.