

Train the Trainers — An Agentic AI Framework for Peer-Based Mental Health Support in Battlefield Environments

Atmaram Yarlagadda^a, Eranga Bandara^b, Ross Gore^b, Anita H. Clayton^d, Preston Samuel^c, Christopher K. Rhea^b, Sachin Shetty^b, Ravi Mukkamala^b, Xueping Liang^e, Amin Hass^f, Abdul Rahman^g

^a*McDonald Army Health Center, Newport News, VA, USA*

^b*Old Dominion University, Norfolk, VA, USA*

^c*Blanchfield Army Community Hospital, Fort Campbell, KY, USA*

^d*Department of Psychiatry and Neurobehavioral Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA*

^e*Florida International University, FL, USA*

^f*Accenture Technology Labs, Arlington, VA, USA*

^g*Deloitte & Touche LLP, USA*

Abstract

Modern military operations expose soldiers to sustained psychological stress, leading to acute reactions, post-traumatic stress symptoms, and other mental health issues. Although the U.S. Department of Defense offers evidence-based therapies, access to trained professionals in forward-deployed and contested environments is limited. As a result, soldiers with early-stage distress are often evacuated to rear medical facilities, delaying care, reducing readiness, and increasing long-term risks. This paper proposes a Train-the-Trainers framework in which soldiers who have completed therapy and returned to duty are trained as peer facilitators to provide first-line psychological support in operational settings. To scale and standardize this

Email addresses: atmaram.yarlagadda.civ@health.mil (Atmaram Yarlagadda), cmedawer@odu.edu (Eranga Bandara), rgore@odu.edu (Ross Gore), AHC8V@uvahealth.org (Anita H. Clayton), preston.l.samuel.mil@health.mil (Preston Samuel), crhea@odu.edu (Christopher K. Rhea), sshetty@odu.edu (Sachin Shetty), mukka@odu.edu (Ravi Mukkamala), xuliang@fiu.edu (Xueping Liang), amin.hassanzadeh@accenture.com (Amin Hass), abdulrahman@deloitte.com (Abdul Rahman)

model under severe resource and connectivity constraints, we introduce an agentic AI-enabled platform that augments these recovered soldiers with specialized AI agents. The recovered soldier acts as a human supervisor, coordinating agents for symptom triage, guided peer-support interventions, operational constraint reasoning, training and simulation, and structured documentation for clinical escalation when needed. To enable responsible and explainable decisions in high-stakes environments, the AI agents use a consortium of fine-tuned language models plus a reasoning-oriented model, applying consensus-based inference to generate recommendations. This approach reduces single-model bias, improves robustness, and provides transparent justifications while preserving human-in-the-loop control. The architecture functions in air-gapped and low-connectivity settings, maintaining human oversight and ethical safeguards. A functional prototype was developed with the McDonald U.S. Army Health Center, Newport News, VA, USA. By combining peer-based intervention with consensus-driven agentic AI decision support, the framework seeks to cut response times, prevent symptom escalation, reduce unnecessary evacuations, and improve continuity of care. This work shows how agentic AI can serve as a force multiplier for mental health support in austere environments and identifies pathways for broader evaluation and deployment across defense and humanitarian operations.

Keywords: Psychiatric Diagnosis, Agentic AI, Responsible AI, Explainable AI, LLM, DSM-5

1. Introduction

Modern military operations increasingly expose soldiers to prolonged psychological stress arising from combat engagement, asymmetric warfare, extended deployments, and morally injurious experiences. These conditions contribute to a wide spectrum of mental health challenges, including acute stress reactions, anxiety disorders, depression, post-traumatic stress symptoms, and cognitive impairment [1]. Early identification and timely intervention are critical, as untreated or delayed psychological distress can escalate into severe long-term conditions, negatively affecting both individual well-being and overall force readiness.

The U.S. Department of Defense (DoD) has established evidence-based therapeutic programs delivered by trained mental health professionals, in-

cluding psychologists, psychiatrists, and specialized therapists. However, the availability of such professionals is inherently limited in forward-deployed and contested environments [2, 3]. Battlefield conditions, logistical constraints, operational tempo, and intermittent or absent connectivity often prevent timely access to professional assessment and care. As a result, soldiers experiencing early or moderate psychological distress are frequently evacuated to rear medical facilities or base hospitals, including those located outside the theater of operations. While clinically necessary in severe cases, such evacuations can delay intervention, disrupt unit cohesion, reduce operational effectiveness, and impose additional logistical and human costs [4].

To address these limitations, military mental health research has increasingly explored peer-based support models, recognizing that soldiers often trust and engage more readily with peers who share similar operational experiences [5]. In this context, soldiers who have successfully undergone therapeutic interventions and returned to duty represent an underutilized resource. Their lived experience and targeted training enable them to recognize early warning signs, provide immediate peer support, and escalate to professional care when needed. This observation motivates a Train-the-Trainers approach, in which recovered soldiers are trained as peer facilitators capable of delivering first-line psychological support within the battlefield environment. Importantly, this model does not seek to replace professional clinicians but rather to augment existing care pathways by enabling early intervention and informed decision-making at the point of need.

Despite its promise, implementing a Train-the-Trainers model at scale presents several challenges. Peer facilitators must operate under high cognitive load, rapidly evolving operational constraints, and limited access to expert guidance [6]. Ensuring consistency, safety, and ethical oversight across peer-led interventions is particularly difficult in austere environments. Without structured support, peer facilitators may struggle to assess symptom severity, select appropriate interventions, or determine when escalation is necessary [2].

Recent advances in agentic artificial intelligence (AI) offer a promising pathway to address these challenges [7, 8]. Agentic AI systems consist of multiple specialized agents capable of perception, reasoning, decision support, and coordination, operating under human supervision. Rather than functioning as autonomous decision-makers, such systems are designed to augment human judgment, provide structured guidance, and manage complex workflows across constrained environments. Prior work has demon-

strated the effectiveness of human-in-the-loop multi-agent architectures in domains requiring high reliability, adaptability, and accountability [9].

In this paper, we propose an agentic AI-enabled Train-the-Trainers framework for battlefield mental health support. In the proposed system, a recovered soldier, trained as a peer facilitator, acts as a human supervisor. This supervisor orchestrates a suite of specialized AI agents that support symptom triage, guided peer interventions, operational constraint reasoning, training and simulation, and structured documentation [10]. The platform is designed to operate in air-gapped or low-connectivity environments, making it suitable for deployment in forward and contested operational settings [11]. Human oversight is preserved at all times, with AI agents providing decision support rather than autonomous clinical judgment. To promote responsible, robust, and explainable decision-making in high-stakes military contexts, the AI agents are configured to operate over a consortium of fine-tuned language models combined with a reasoning-oriented model, using a consensus-driven inference mechanism to generate recommendations [12, 13]. This approach mitigates single-model bias, improves reliability under uncertainty, and enables transparent justification of agent outputs, while maintaining strict human-in-the-loop control and supporting responsible and explainable AI principles [14]. By positioning recovered soldiers as peer facilitators, supported by consensus-driven agentic AI, the proposed framework leverages both lived experience and structured decision support. This combined approach addresses the challenges of timely assessment, peer-level intervention, and informed escalation in austere and resource-constrained environments. The primary contributions of this paper are fourfold:

1. The formulation of a Train-the-Trainers model tailored to battlefield mental health support
2. The design of an agentic AI architecture that augments recovered soldiers as peer facilitators under operational constraints
3. A conceptual workflow illustrating how human-AI collaboration can reduce response latency, prevent symptom escalation, and improve continuity of care.
4. The implementation of a functional prototype of the proposed platform, developed in collaboration with the McDonald U.S. Army Health Center, Newport News, VA, USA.

The remainder of this paper is organized as follows. Section 2 provides background on the key concepts underpinning this work, including large lan-

guage models, agentic AI, and responsible and explainable AI. Section 3 describes the entities involved in the proposed Train-the-Trainers framework, including human roles and specialized AI agents. Section 4 details the operational functionality of the system, focusing on the peer-support workflow and human–AI interaction model. Section 5 presents the implementation and evaluation of the proposed platform. Section 6 reviews related work in military mental health support, peer-based interventions, and agentic AI systems. Finally, Section 7 concludes the paper and outlines directions for future research.

2. Background

This section introduces the key technical concepts underlying the proposed Train-the-Trainers framework. We briefly review large language models, reasoning-oriented language models, fine-tuning approaches, agentic AI systems, and responsible and explainable AI principles, which together form the foundation of the proposed architecture.

2.1. Large Language Models (LLMs)

Large Language Models (LLMs) are deep neural networks trained on large-scale textual corpora to learn statistical representations of language [15]. Through pre-training on diverse data, LLMs acquire the ability to generate coherent text, perform question answering, summarize information, and reason over complex natural language inputs. Recent advances have demonstrated that LLMs can be adapted to domain-specific tasks through fine-tuning, enabling their use in decision-support applications across health-care, defense, and other high-stakes domains [16].

In the context of this work, LLMs serve as the core reasoning engines that process structured and unstructured assessment inputs, generate interpretable outputs, and support human decision-making rather than autonomous action.

2.2. Reasoning-Oriented Language Models

Reasoning-oriented language models extend standard LLM capabilities by emphasizing multi-step inference, comparison of alternatives, and explicit justification of outputs. These models are designed to evaluate multiple candidate responses, assess coherence and contextual relevance, and synthesize a

final output that reflects structured reasoning rather than pattern completion alone [17, 18].

Within the proposed framework, a reasoning-oriented LLM acts as a supervisory component that synthesizes outputs from multiple fine-tuned LLMs. This consensus-driven reasoning approach improves robustness, reduces single-model bias, and enables the presentation of transparent rationales to support explainable and accountable decision support.

2.3. LLM Fine-Tuning

Fine-tuning is the process of adapting a pre-trained LLM to a specific task or domain by training it further on curated datasets. Techniques such as Low-Rank Adapters (LoRA) and quantization-aware fine-tuning enable efficient adaptation while reducing computational and memory requirements, making fine-tuned models suitable for deployment in resource-constrained environments [19, 3].

In this work, fine-tuning is used to specialize LLMs for structured psychological assessment tasks. Fine-tuning is performed offline or through secure update pipelines, ensuring that system behavior remains stable during live operation and compatible with air-gapped deployment requirements.

2.4. AI Agents and Agentic AI

AI agents represent a paradigm in which LLMs are embedded within autonomous or semi-autonomous software entities capable of interacting with their environment, invoking tools, maintaining state, and executing multi-step workflows. An AI agent can autonomously perform an interaction loop by constructing prompts, invoking language models, interpreting responses, and triggering subsequent actions without direct human intervention [7, 9].

Figure 1 contrasts traditional human–LLM interaction with autonomous AI agent–LLM interaction. In essence, AI agents are software programs that leverage LLMs in combination with tools, application programming interfaces (APIs), and external context to execute tasks iteratively. Agentic AI systems extend this concept by coordinating multiple specialized agents under a shared orchestration framework.

In the proposed Train-the-Trainers framework, agentic AI is employed in a human-supervised configuration, where AI agents provide structured decision support while authority and execution remain under human control.

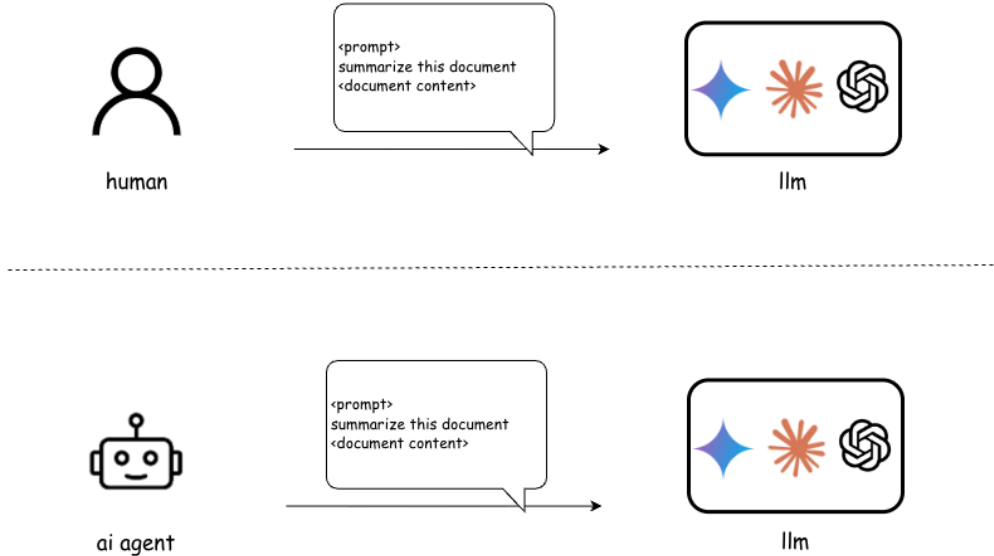


Figure 1: Comparison between direct human–LLM interaction and agentic AI–LLM interaction. In the top panel, a human user directly constructs and submits a prompt to a language model, manually interpreting the response. In the bottom panel, an AI agent autonomously manages the interaction loop by constructing prompts, invoking the language model, interpreting responses, and determining subsequent actions without direct human intervention. This agent-mediated workflow enables iterative reasoning, tool integration, and task automation, forming the basis of agentic AI systems used for structured decision support and multi-step workflows.

2.5. Responsible and Explainable AI

Responsible and Explainable AI principles emphasize transparency, accountability, fairness, and human oversight in AI system design. In high-stakes domains such as military operations and mental health support, these principles are critical for ensuring trust, safety, and ethical deployment [14].

Explainable AI focuses on providing interpretable outputs and rationales that allow human operators to understand and evaluate system recommendations. Responsible AI further requires clear operational boundaries, avoidance of autonomous decision-making in sensitive contexts, and mechanisms for auditability and oversight [20, 21].

The proposed framework operationalizes these principles through strict human-in-the-loop control, consensus-driven reasoning across multiple models, structured logging of agent interactions, and offline learning mechanisms.

This design ensures that agentic AI augments human capability without displacing human judgment or authority.

3. Entities in the System

This section describes the primary human and artificial entities involved in the proposed Train-the-Trainers framework. The system is designed to preserve clear boundaries of authority, ensuring that all decision-making remains under human supervision while AI agents provide structured decision support in austere and resource-constrained environments.

3.1. Human Entities

3.1.1. Peer Trainer (*Recovered Soldier*)

The peer trainer is a recovered soldier who has successfully completed therapeutic intervention and returned to active duty. In the proposed framework, the peer trainer serves as the *human supervisor* of the agentic AI platform and is responsible for initiating, overseeing, and executing all interactions within the system [10]. The peer trainer is not a licensed clinician and does not perform diagnosis or treatment; instead, they provide first-line peer support informed by training and augmented by AI-based decision support.

The peer trainer retains full authority over all actions, including whether to engage AI agents, how to interpret their recommendations, and when to escalate a situation to professional clinical services. By combining lived operational experience with structured guidance from AI agents, the peer trainer is positioned to deliver timely and context-aware support in forward-deployed environments while maintaining ethical and procedural boundaries [22].

3.1.2. Soldier Receiving Support

The soldier receiving support is an active-duty service member experiencing psychological distress during deployment or operational activities. Interactions between the peer trainer and the supported soldier occur within a peer-to-peer context, emphasizing trust, shared experience, and reduced stigma. Participation is voluntary, and the system is designed to respect privacy while minimizing unnecessary data collection [5].

The supported soldier does not interact directly with the AI agents. All AI-mediated guidance is filtered through the peer trainer, ensuring that technology augments human interaction rather than replacing it. This design reinforces human-centered care and reduces the risk of inappropriate reliance on automated systems [9].

3.2. AI Agents

The agentic AI platform comprises multiple specialized AI agents, each responsible for a narrowly scoped function. All agents operate under strict human-in-the-loop control and do not perform autonomous clinical decision-making, as illustrated in Figure 2. Recommendations are generated using a consensus-driven inference process over a consortium of fine-tuned language models in combination with a reasoning-oriented model, supporting robustness, transparency, and explainability in high-stakes operational environments [14, 23].

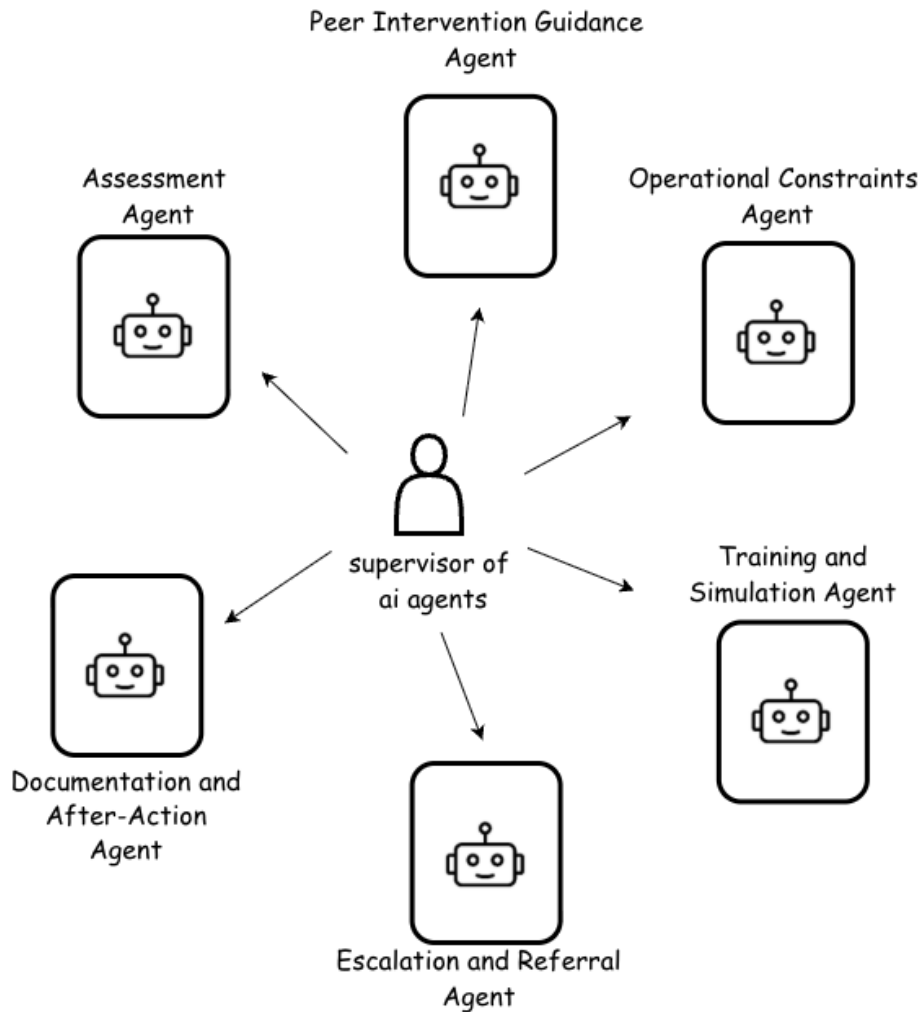


Figure 2: Human-supervised agentic AI architecture for the Train-the-Trainers framework. A recovered soldier acting as a peer trainer serves as the central human supervisor, coordinating multiple specialized AI agents that provide advisory decision support for assessment, peer intervention guidance, operational constraint reasoning, escalation and referral, documentation and after-action reporting, and training and simulation. All agent interactions are advisory in nature, with authority and execution remaining strictly under human control.

3.2.1. Assessment Agent

The Assessment Agent supports structured, non-clinical evaluation of observable psychological and behavioral indicators during peer-support inter-

actions. It assists the peer trainer in systematically checking symptoms, contextual signals, and functional impairments without performing a medical diagnosis. The agent operationalizes an assessment framework that captures early indicators of psychological distress, such as acute stress responses, emotional dysregulation, cognitive overload, sleep disruption, and behavioral withdrawal [24].

The assessment process is guided by a structured framework that we previously developed for systematic symptom and context evaluation. Within the Train-the-Trainers system, this framework is embedded as an assessment logic layer that translates observed signals into interpretable assessment outputs, including severity bands, risk flags, and uncertainty indicators. The agent presents these outputs to the peer trainer as decision-support artifacts, enabling consistent and repeatable assessment across different operational contexts while explicitly avoiding clinical labeling or diagnosis [25, 26].

By standardizing early assessment, the Assessment Agent reduces subjective variability, supports early identification of concerning patterns, and provides a transparent basis for subsequent intervention or escalation decisions.

3.2.2. Peer Intervention Guidance Agent

The Peer Intervention Guidance Agent provides real-time decision support during peer-led interactions following assessment. Based on assessment outputs and ongoing observations, the agent suggests appropriate non-clinical stabilization and support techniques, including grounding exercises, breathing regulation, short rest-cycle interventions, and structured peer communication strategies [27, 28].

Recommendations adapt dynamically as the situation evolves, but remain strictly advisory. The agent does not prescribe actions or override human judgment; instead, it supports the peer trainer by reducing cognitive load during high-stress interactions and ensuring alignment with established peer-support practices.

3.2.3. Operational Constraints Agent

The Operational Constraints Agent reasons over contextual factors specific to battlefield and forward-deployed environments. These factors include mission posture, location, time sensitivity, unit availability, communication constraints, and access to medical or evacuation resources [29].

By integrating operational context with assessment and intervention in-

formation, the agent assists the peer trainer in identifying feasible courses of action that balance mental health support needs with mission requirements. This ensures that recommended interventions and escalation pathways remain practical, actionable, and aligned with operational realities [30].

3.2.4. Escalation and Referral Agent

The Escalation and Referral Agent supports decision-making when peer-level intervention is insufficient or when assessment outputs indicate elevated risk. It assists the peer trainer in recognizing escalation thresholds and identifying appropriate referral pathways, such as on-site medical personnel, remote consultation when available, or evacuation to rear medical facilities [24].

The agent generates structured handoff summaries that consolidate assessment findings, observed behaviors, actions taken, and contextual constraints. These summaries are designed to improve continuity of care and reduce information loss during transitions to professional services. Final escalation decisions remain entirely under the authority of the peer trainer.

3.2.5. Documentation and After-Action Agent

The Documentation and After-Action Agent assists in capturing minimal, structured records of peer-support interactions. It summarizes key assessment outputs, interventions applied, decisions made, and observed outcomes in a standardized format suitable for later review by authorized personnel [26].

The agent is designed to minimize administrative burden and cognitive overhead while supporting accountability, continuity of care, and after-action analysis. All documentation is privacy-preserving and avoids unnecessary personal or identifying information.

3.2.6. Training and Simulation Agent

The Training and Simulation Agent supports both initial preparation and ongoing skill maintenance for peer trainers. It provides scenario-based training exercises and simulated interactions that reflect common battle-field psychological stress situations, including ambiguous, time-pressured, and resource-constrained cases [31].

Using feedback from assessment, intervention, and escalation outcomes, the agent offers structured feedback to peer trainers on communication effectiveness, assessment consistency, and decision-making quality. This enables

continuous skill development without requiring constant access to human instructors [32].

3.2.7. Fine-Tuning and Continuous Learning Agent

The Fine-Tuning and Continuous Learning Agent enables adaptive improvement of the agentic AI platform over time. It aggregates inputs and outputs generated by other AI agents, including assessment results, intervention guidance, operational reasoning, escalation summaries, and training interactions, along with human feedback provided by peer trainers. These data are processed in a privacy-preserving and non-identifying manner [19].

Based on the aggregated information, the agent supports offline or securely managed fine-tuning of the language models that form the system’s model consortium. Updated models are incorporated into subsequent system iterations, enabling a continuous learning approach that improves robustness, contextual relevance, and decision-support quality while preserving human oversight and system stability [33, 34].

4. Functionality of the System

This section describes how the proposed Train-the-Trainers framework operates in practice. Rather than reiterating individual agent roles, the focus here is on the end-to-end operational workflow, the lifecycle of peer trainers, and the manner in which agentic AI supports human decision-making throughout training, deployment, and peer-support interactions in battlefield environments.

4.1. Lifecycle of Peer Trainers

Soldiers experiencing psychological distress during deployment are initially removed from active battlefield duties and referred to rear medical facilities or base hospitals, where they receive appropriate professional therapeutic interventions conducted by qualified mental health professionals. These interventions follow established clinical protocols and aim to support recovery, stabilization, and functional readiness [6].

Upon successful completion of therapy and clearance for return to duty, a subset of recovered soldiers may voluntarily enroll in additional preparation programs to serve as peer trainers. These programs focus on recognizing early indicators of psychological distress, applying non-clinical peer-support techniques, understanding escalation pathways, and operating within defined

ethical and procedural boundaries. Scenario-based exercises and simulated interactions are used to reinforce decision-making and communication skills.

Soldiers who complete these requirements graduate as peer trainers and are equipped with the agentic AI-enabled platform prior to redeployment.

4.2. Peer Support Interaction Workflow in Battlefield Environments

Once deployed, peer trainers return to forward-deployed or contested operational environments as part of their assigned units. In addition to their primary operational responsibilities, they serve as first-line peer-support facilitators within their units. When a fellow soldier exhibits signs of psychological distress, the peer trainer initiates a peer-support interaction and engages the agentic AI platform as needed [5].

The peer-support workflow proceeds through a sequence of stages. First, the peer trainer conducts an initial engagement with the supported soldier, relying on direct observation and peer communication. The system then provides structured assessment outputs that help the peer trainer evaluate observable symptoms, contextual factors, and functional impacts in a consistent and non-clinical manner [4]. Based on these assessment outputs, the peer trainer may apply appropriate peer-level stabilization and support techniques, such as grounding, breathing regulation, or short rest-cycle interventions.

Throughout the interaction, the system incorporates operational context, including mission posture, time constraints, and available resources, to ensure that recommended actions remain feasible in the battlefield environment. At defined decision points, the peer trainer determines whether peer-level intervention is sufficient or whether escalation to professional medical services is required. If escalation is necessary, the system supports structured handoff by consolidating relevant observations and actions into a concise summary for medical personnel [2].

All interactions remain human-centered: the peer trainer is the sole point of contact with the supported soldier, and all system outputs are advisory and subject to human judgment.

4.3. Human-AI Interaction Model

The proposed framework adopts a strict human-in-the-loop interaction model. The agentic AI platform does not act autonomously or initiate actions independently. Instead, it provides structured decision support at spe-

cific stages of the peer-support workflow, presenting recommendations, risk indicators, and contextual considerations to the peer trainer [35, 23].

In cases where multiple perspectives are available, system outputs are derived through a consensus-driven reasoning process that synthesizes inputs from multiple fine-tuned language models and a reasoning-oriented model. Recommendations are accompanied by interpretable rationales to support transparency and informed human decision-making. The peer trainer retains full authority to accept, modify, or disregard any system recommendation [10].

4.4. Continuous Learning and System Evolution

System learning occurs outside of live peer-support interactions to preserve stability, safety, and operational reliability. Following deployments or structured training cycles, the system aggregates structured outputs generated during assessments, peer-level interventions, escalation decisions, documentation activities, and training exercises, together with human feedback provided by peer trainers. All collected data are processed in a privacy-preserving and non-identifying manner [19, 36, 37].

The aggregated information supports offline or securely managed fine-tuning of the language models that form the system’s model consortium. Fine-tuning is performed between operational cycles and does not modify system behavior during active use. Updated models are incorporated into subsequent system iterations, enabling continuous improvement in contextual awareness, robustness, and decision-support quality while maintaining compatibility with air-gapped and low-connectivity environments. Figure 3 illustrates the fine-tuning workflow and deployment process used to update language models safely across iterations [38, 33].

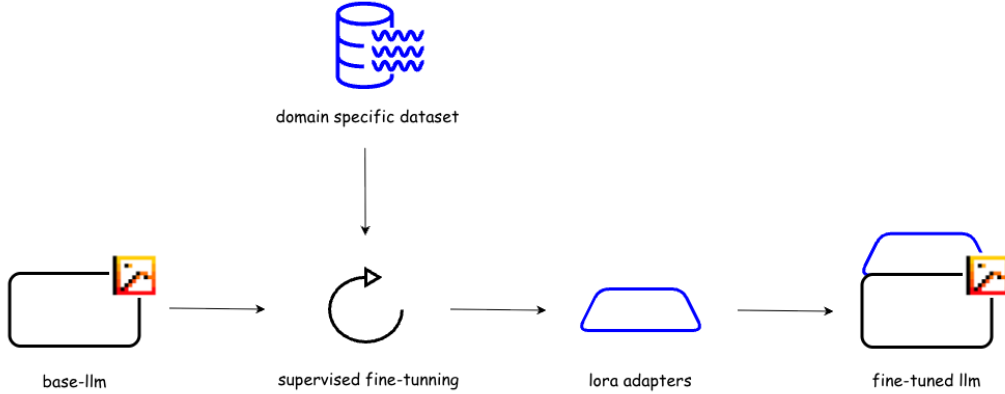


Figure 3: Overview of the supervised fine-tuning pipeline used to adapt a base LLM to domain-specific tasks. A curated, domain-specific dataset is used to perform supervised fine-tuning of a base LLM, with Low-Rank Adapter (LoRA) modules applied to enable parameter-efficient training. The resulting fine-tuned model retains the general linguistic capabilities of the base model while incorporating domain-specific knowledge, supporting efficient deployment and updates in resource-constrained and air-gapped environments.

4.5. Operational Boundaries and Safety Considerations

The proposed system is explicitly designed as a decision-support platform and does not perform clinical diagnosis, treatment, or autonomous intervention. All recommendations generated by the agentic AI platform are advisory in nature, and all interpretation, decision-making, and execution remain under direct human control [8, 23]. The peer trainer serves as the sole authority responsible for actions taken during peer-support interactions.

To support responsible and explainable AI use in high-stakes environments, the system employs a consensus-driven reasoning architecture in which recommendations are synthesized across multiple fine-tuned language models and a reasoning-oriented model. This design promotes robustness, reduces single-model bias, and enables the presentation of interpretable rationales alongside recommendations [14]. The system operates locally without reliance on continuous network connectivity, making it suitable for deployment in austere and contested environments. Figure 4 illustrates the consensus-driven LLM consortium and reasoning flow that underpin responsible, transparent, and human-supervised decision support.

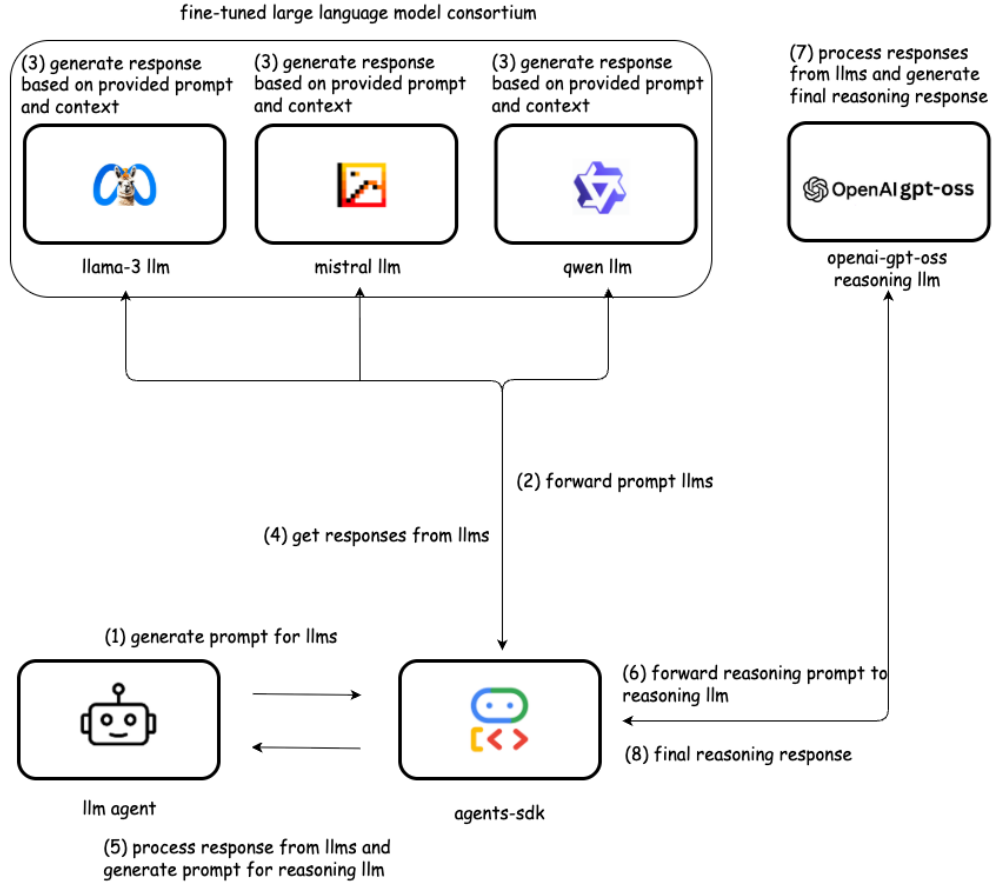


Figure 4: Consensus-driven LLM consortium and reasoning workflow used by the agentic AI platform. An LLM agent generates a task-specific prompt, which is forwarded via the agent orchestration layer to multiple fine-tuned LLMs (e.g., Llama-3, Mistral, and Qwen). Each model independently generates a response based on the same prompt and contextual inputs. The resulting outputs are collected and processed by the orchestration layer, which constructs a structured reasoning prompt for a dedicated reasoning-oriented LLM (GPT-OSS). The reasoning LLM evaluates, reconciles, and synthesizes the multiple model outputs to produce a final consolidated response accompanied by an interpretable rationale. This consensus-driven process improves robustness, mitigates single-model bias, and supports explainable, human-supervised decision support.

5. Implementation and Evaluation

This section presents the implementation and evaluation of the proposed Train-the-Trainers platform. We describe the realization of a functional

prototype designed to demonstrate the feasibility of deploying a human-supervised, agentic AI system for peer-based mental health support in forward-deployed and resource-constrained environments [10]. The section then outlines the primary assessment use case implemented within the platform and defines the scope of the evaluation conducted to assess the technical performance of the underlying agentic AI components [39].

5.1. Implementation Overview

The Train-the-Trainers framework was implemented as a functional prototype to show the feasibility of a human-supervised, agentic AI platform for peer-based mental health support in forward-deployed, resource-constrained settings. It supports recovered soldiers serving as peer trainers by enabling structured assessment, intervention guidance, escalation, documentation, training and simulation, and continuous learning without requiring constant network connectivity.

The system integrates multiple specialized AI agents, each responsible for a distinct support function, including assessment, peer intervention guidance, operational constraint reasoning, escalation and referral support, documentation and after-action reporting, training and simulation, and fine-tuning with continuous learning [3]. All agents operate under a strict human-in-the-loop orchestration model, with the peer trainer serving as the sole authority responsible for interpreting recommendations and executing actions.

Each AI agent was implemented using the OpenAI Agents SDK [7, 40], which provides modular primitives for defining agent roles, behaviors, and reasoning strategies. To enable interoperability and structured agent-to-agent communication, all agent functionalities were exposed through a Model Context Protocol (MCP) [41, 42, 43] server interface. This design allows agents to exchange structured assessment outputs, contextual signals, and recommendation artifacts through standardized endpoints while maintaining clear boundaries of responsibility.

For human interaction, the prototype employs an MCP-compatible interface tool that allows peer trainers to interact with individual agents, review consensus-driven recommendations, and control system behavior without requiring direct code-level intervention [23, 9]. This interface supports transparent orchestration, enabling peer trainers to supervise multi-agent interactions, request specific support functions, and review structured rationales accompanying agent outputs.

To support deployment in contested and austere environments, the prototype was designed to operate in air-gapped or low-connectivity settings [44]. All core functionalities, including agent execution, reasoning, and decision support, run locally. Network connectivity is required only for periodic system updates or offline fine-tuning workflows, which are performed outside active operational contexts.

LLM fine-tuning was conducted using the Unsloth library [45] on GPU-enabled environments equipped with NVIDIA A100 accelerators [46, 47]. To ensure efficient training and compatibility with constrained hardware, Low-Rank Adapters (LoRA) [37] combined with 4-bit quantization (QLoRA) [48, 49] were employed. The system utilizes a consortium of open-source language models, including Llama-3, Pixtral, and Qwen, each fine-tuned for specific support functions such as structured assessment, communication guidance, and operational reasoning [50, 51, 52].

A dedicated reasoning-oriented language model (GPT-OSS) serves as the supervisory component within the consortium, responsible for synthesizing and validating outputs generated by individual models [53, 17, 18]. This consensus-driven reasoning process improves robustness, reduces single-model bias, and enables the generation of interpretable rationales to support explainable and responsible decision support.

Overall, the prototype demonstrates that the proposed agentic AI architecture can be effectively implemented using MCP-based orchestration and fine-tuned language models, operating reliably in air-gapped and resource-constrained environments [43]. These results indicate the feasibility of deploying the Train-the-Trainers framework as a scalable and adaptable decision-support system for peer-based mental health support in military and humanitarian contexts.

5.2. Assessment Agent Use Case and Workflow

The primary use case evaluated in this work is a psychiatric assessment support workflow implemented as the Assessment Agent within the agentic AI platform. The Assessment Agent is designed to support structured, non-clinical evaluation of observable psychological and behavioral indicators, enabling peer trainers to perform consistent early assessment without performing medical diagnosis.

The workflow follows a consensus-driven reasoning approach. Structured assessment inputs derived from observable symptoms, contextual information, and interaction cues are processed by a consortium of fine-tuned LLMs,

each specialized to capture different aspects of assessment reasoning [14]. These models generate independent assessment outputs reflecting diverse reasoning perspectives. A reasoning-oriented LLM (GPT-OSS) then evaluates, compares, and synthesizes these outputs to produce a consolidated assessment representation accompanied by an interpretable rationale.

This design aligns with the assessment framework discussed in related work and reflects realistic operational usage, where peer trainers rely on structured decision support rather than autonomous classification or diagnosis. The Assessment Agent operates entirely in an advisory capacity, with all interpretations and decisions remaining under human control.

5.3. Evaluation of LLM Fine-Tuning

To evaluate the fine-tuning process used to support the Assessment Agent, we fine-tuned the Qwen2.5-7B-Instruct model [52] on a curated psychiatric assessment dataset. The dataset consists of 500 synthetic psychiatrist–patient conversations spanning five DSM-5 reference categories: Major Depressive Disorder (DSM-5 296.2x), Generalized Anxiety Disorder (DSM-5 300.02), Panic Disorder (DSM-5 300.01), Post-Traumatic Stress Disorder (DSM-5 309.81), and Bipolar I Disorder (DSM-5 296.4x) [54, 55]. These categories were used as structured reference labels to evaluate assessment alignment rather than to enable autonomous diagnosis.

The fine-tuning dataset was constructed using a combination of publicly available psychiatric dialogue data and synthetically generated conversations. Synthetic samples were generated using large language models guided by established diagnostic criteria to ensure realistic symptom expression, contextual coherence, and adherence to DSM-5 assessment patterns [56, 31]. Figure 5 illustrates the structure of the fine-tuning dataset.

Search this dataset

instruction string · classes	conversation string · lengths	diagnosis string · classes	condition string · classes
1 value	 2.07k 5.07k	5 values	5 values
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I've been feeling stuck in this weird low mood for a while, like I'm_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I'm 44, and for about six weeks I've felt_ off. I can't shake this flat,_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I'm not sure what's wrong with me, but for the last few months I've fel_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I've been feeling really off for the past month or two. I can't shake thi_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I'm here because for the past couple of months I just haven't felt lik_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I'm sorry if I ramble. I just_ I haven't felt like myself for a while, an_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I've been feeling off for like a month and a half now. It's not a crisis_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I don't know what's wrong with me, but the last couple of months I've_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I'm here because I just can't seem to pull myself out of this hole. Fo_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I've just been feeling off for a while-kind of heavy and stuck. I'm_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I've been dragging myself through the days for weeks now. My wife_	Major Depressive Disorder (DSM-5 296.2x)	depression
You are a psychiatric diagnostic assistant. Analyze the following_	Patient: I came in because I can't seem to pull myself together anymore. For the_	Major Depressive Disorder (DSM-5 296.2x)	depression

< Previous 1 2 3 ... 5 Next >

Figure 5: Structure of the psychiatric assessment fine-tuning dataset. Each sample contains an instruction prompt, a psychiatrist-patient conversation (2.07k–5.07k characters), a DSM-5 diagnosis label, and a condition category. The dataset was constructed from publicly available data and synthetically generated conversations covering five psychiatric conditions.

Each training sample consists of four components: (1) *instruction*, a standardized system prompt directing the model to perform structured assessment reasoning; (2) *conversation*, a full psychiatrist–patient dialogue ranging from 2,070 to 5,070 characters and capturing symptom presentation, patient history, and assessment-oriented questioning; (3) *diagnosis*, representing the reference DSM-5 category used for supervised learning; and (4) *condition*, a categorical label identifying the corresponding assessment class. Table 1 summarizes the dataset composition and class distribution [31].

Table 1: Psychiatric Assessment Dataset Statistics

Category	Count	Percentage
Total Samples	500	100.0%
Training Set	400	80.0%
Validation Set	50	10.0%
Test Set	50	10.0%
<i>Training Set Distribution by Condition:</i>		
Depression	82	20.5%
Panic Disorder	83	20.8%
PTSD	83	20.8%
Bipolar Disorder	79	19.8%
Anxiety Disorder	73	18.3%

Fine-tuning was performed using the Unsloth library on Google Colab with an NVIDIA A100-SXM4-40GB GPU [45, 46]. The base model was loaded in 4-bit quantized format to optimize memory utilization. We employed Low-Rank Adaptation (LoRA) [37] with rank $r = 32$ and $\alpha = 32$, targeting the attention and MLP projection layers (q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj). This configuration resulted in 40,370,176 trainable parameters, representing only 0.53% of the total 7.66 billion model parameters. Table 2 details the complete training configuration.

Table 2: Fine-Tuning Configuration for Psychiatric Assessment Model

Parameter	Value
<i>Model Configuration:</i>	
Base Model	Qwen2.5-7B-Instruct (4-bit)
Maximum Sequence Length	4,096 tokens
Precision	BFloat16
<i>Training Configuration:</i>	
Per-device Batch Size	2
Gradient Accumulation Steps	4
Effective Batch Size	8
Maximum Training Steps	200 (≈ 4 epochs)
Learning Rate	1×10^{-4}
Warmup Steps	20
Learning Rate Scheduler	Linear decay
Optimizer	AdamW (8-bit)
Weight Decay	0.01
Early Stopping Patience	10 evaluations
<i>LoRA Configuration:</i>	
LoRA Rank (r)	32
LoRA Alpha (α)	32
LoRA Dropout	0
Target Modules	q, k, v, o, gate, up, down proj
Trainable Parameters	40,370,176 (0.53%)
Total Model Parameters	7,655,986,688

5.3.1. Training Convergence Analysis

Training and validation loss curves were monitored via TensorBoard to assess convergence behavior and detect potential overfitting. Figure 6 illustrates the training loss progression, which decreased from an initial value of approximately 1.9 to 0.90 over 180 training steps, demonstrating consistent learning without significant instability.

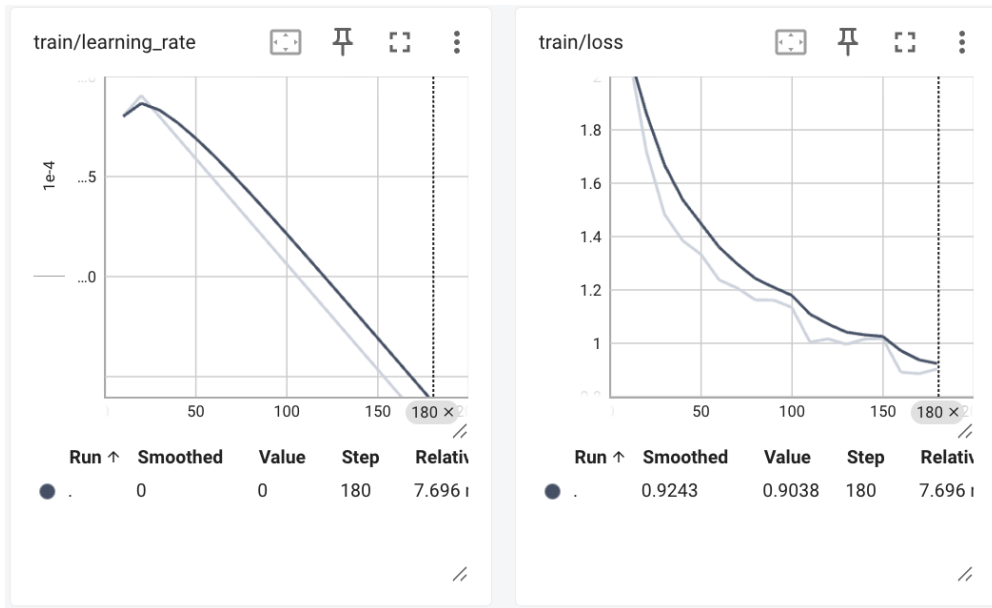


Figure 6: Training metrics during fine-tuning: learning rate schedule (left) showing warmup to 1×10^{-4} followed by linear decay, and training loss (right) decreasing from 1.9 to 0.90 over 180 steps, indicating effective learning of the diagnostic classification task.

The validation loss, shown in Figure 7, decreased from approximately 1.8 to 1.195, closely tracking the training loss without significant divergence. This parallel decrease indicates that the model generalized well to unseen data without overfitting to the training set.

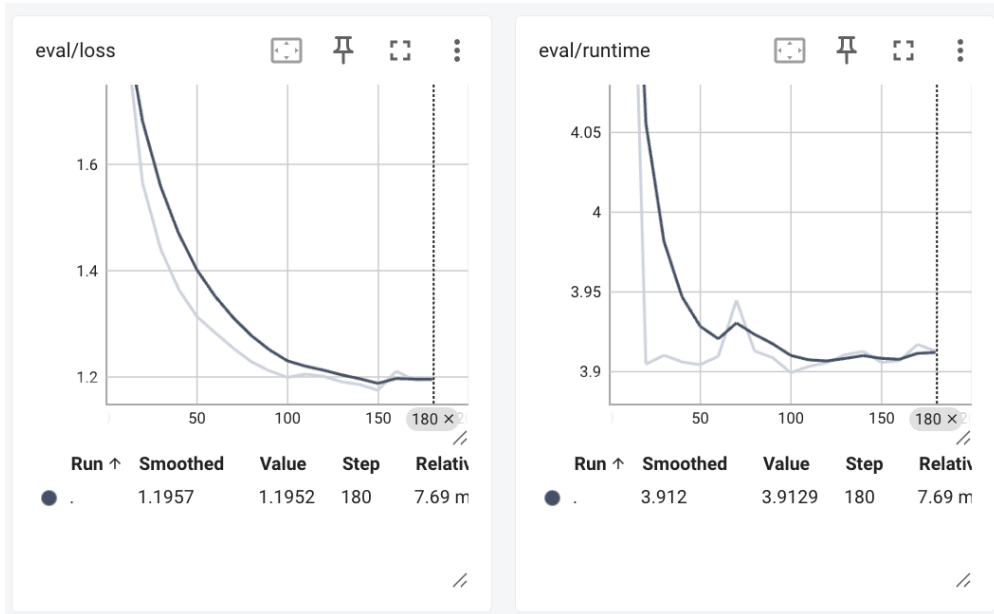


Figure 7: Validation metrics during fine-tuning: evaluation loss (left) decreasing from 1.8 to 1.195, closely following the training loss trajectory and indicating good generalization without overfitting, and evaluation runtime (right) stabilizing at approximately 3.91 seconds.

The gradient norm, depicted in Figure 8, increased gradually from approximately 0.28 to 0.71 throughout training. This controlled increase indicates stable gradient flow without exploding gradients, validating the effectiveness of the chosen hyperparameters and LoRA configuration [36].

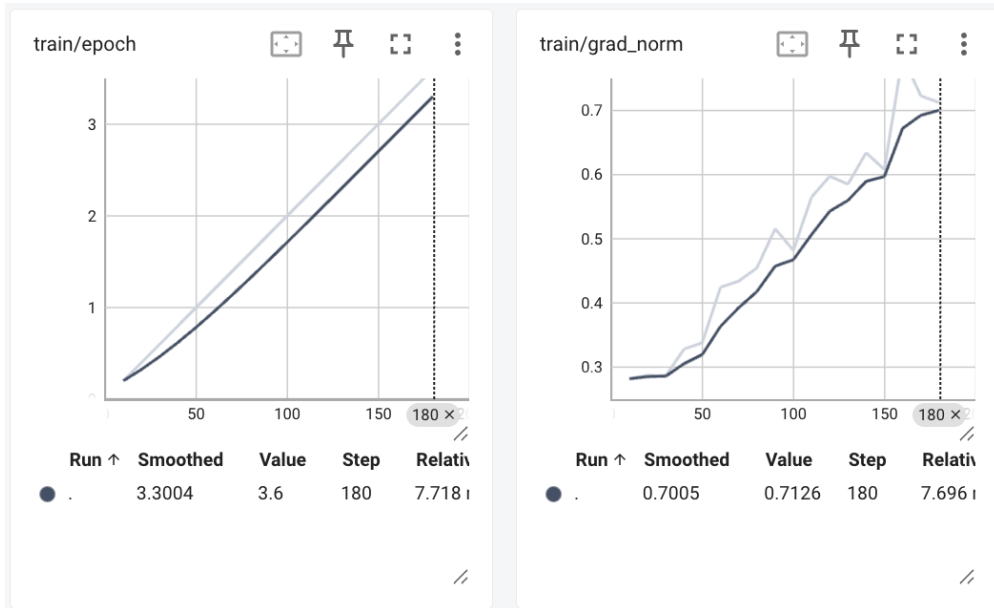


Figure 8: Training progression metrics: epoch count (left) showing linear progression through approximately 3.6 epochs, and gradient norm (right) increasing gradually from 0.28 to 0.71, indicating stable optimization without gradient explosion.

5.3.2. Resource Utilization

Training was completed in 494.24 seconds (8.24 minutes) on a single NVIDIA A100-SXM4-40GB GPU. Table 3 summarizes the computational resource utilization during fine-tuning.

Table 3: Training Resource Utilization for Psychiatric Assessment Model

Metric	Value
Training Duration	494.24 seconds (8.24 minutes)
Hardware	NVIDIA A100-SXM4-40GB
GPU Maximum Memory	39.557 GB
Initial Memory Reserved	6.725 GB
Peak Memory Reserved	7.789 GB
Memory for LoRA Training	1.064 GB
Training Steps Completed	180 (early stopping)
Number of Epochs	≈3.6
Total FLOPs	5.65×10^{16}

The peak reserved memory during training was 7.789 GB, with the initial model loading consuming 6.725 GB. The additional memory required for LoRA fine-tuning was only 1.064 GB, representing approximately 13.7% overhead.

Figure 9 illustrates the total floating-point operations (FLOPs) and cumulative training loss. The fine-tuning process required approximately 5.65×10^{16} FLOPs, demonstrating the computational efficiency of LoRA-based adaptation compared to full model fine-tuning, which would require significantly higher computational resources [37, 38].

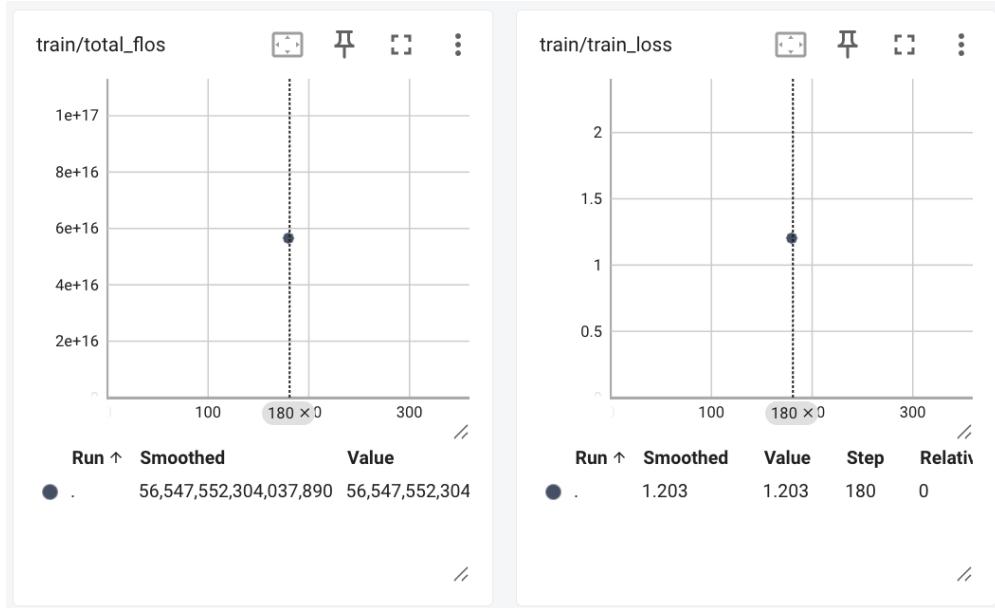


Figure 9: Computational efficiency metrics: total floating-point operations (left) reaching 5.65×10^{16} FLOPs, and cumulative training loss (right) at 1.203, demonstrating the computational cost of parameter-efficient fine-tuning.

These efficiency metrics demonstrate that parameter-efficient fine-tuning enables adaptation of large language models for specialized psychiatric assessment tasks using accessible cloud-based infrastructure.

5.3.3. Inference Performance

Inference throughput metrics, shown in Figure 10, indicate stable evaluation performance throughout training. The model achieved approximately 12.78 samples per second and 6.39 steps per second during validation, with

evaluation runtime stabilizing at approximately 3.91 seconds per evaluation batch.

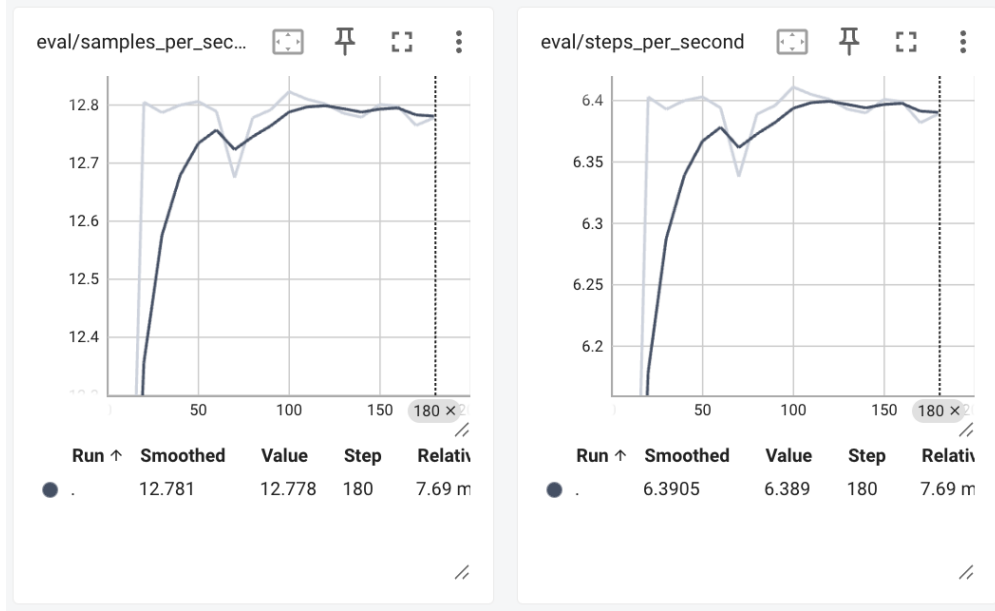


Figure 10: Inference performance metrics during validation: samples per second (left) stabilizing at approximately 12.78, and steps per second (right) at approximately 6.39, indicating consistent evaluation throughput throughout training.

5.4. Predictive Performance of Fine-Tuned LLMs

Following the fine-tuning phase, we evaluated the predictive performance of the fine-tuned LLM consortium in the context of structured psychological assessment support. The evaluation compared outputs generated by baseline (pre-trained) LLMs and their fine-tuned counterparts against reference annotations derived from psychiatrist–patient conversation data. These annotations, including DSM-5 diagnostic codes, were used solely as structured labels for assessing alignment and consistency, rather than as indicators of autonomous diagnostic capability [55].

Quantitative Assessment Alignment. To assess predictive alignment quantitatively, the fine-tuned model was evaluated on a held-out test set comprising 50 assessment samples. Table 4 summarizes the alignment accuracy across five psychiatric reference categories.

Table 4: Assessment Label Alignment Accuracy by Condition

Condition	Correct	Total	Accuracy
Generalized Anxiety Disorder	13	13	100.0%
Bipolar I Disorder	14	14	100.0%
Panic Disorder	8	8	100.0%
Major Depressive Disorder	5	8	62.5%
Post-Traumatic Stress Disorder	4	7	57.1%
Overall	44	50	88.0%

The model achieved perfect alignment for Generalized Anxiety Disorder, Bipolar I Disorder, and Panic Disorder, indicating strong pattern recognition for these symptom clusters. In contrast, performance on Major Depressive Disorder and Post-Traumatic Stress Disorder was limited, reflecting the higher degree of symptom overlap, contextual dependence, and presentation variability associated with these conditions. These results highlight both the potential and limitations of LLM-based assessment support and reinforce the importance of human-in-the-loop oversight within the Train-the-Trainers framework.

Qualitative Assessment Outputs. In addition to quantitative evaluation, we examined representative assessment outputs to assess improvements in structure, clarity, and interpretability following fine-tuning.

Figure 11 illustrates a representative assessment output generated by the fine-tuned Llama-3 model. Prior to fine-tuning, Llama-3 produced verbose and loosely structured responses that referenced relevant symptoms but did not consistently align observations with standardized assessment categories. After fine-tuning, the model generated concise and structured assessment summaries that aligned more reliably with the corresponding reference labels, reflecting improved reasoning over symptom narratives and contextual cues [50].

Conversation	Patient: I've been feeling really down for weeks and I don't enjoy anything anymore. Doctor: How is your sleep and appetite? Patient: I can't sleep and I've lost my appetite.
Real Observation	Major Depressive Disorder (DSM-5 296.21)
Llama-3 Prediction (Before Fine-tuning)	<p>**Major Depressive Disorder (MDD)**</p> <p>Criterion A: Five or more symptoms during the same two-week period, at least one of which is (1) depressed mood most of the day, nearly every day; and Criterion B: The symptoms cause clinically significant distress or impairment in social, occupational, or other areas of functioning.</p> <p>Given the patient's reports of feeling "really down" for weeks, lack of interest in activities, insomnia, and loss of appetite, I believe MDD is a strong contender. The patient's symptoms align with those listed in the DSM-5 criteria for MDD:</p>
Llama-3 Prediction (After Fine-tuning)	Major Depressive Disorder - DSM-5 296.21

Figure 11: Representative assessment output of the fine-tuned Llama-3 LLM aligned with structured reference labels (DSM-5 296.21).

Figure 12 presents a representative assessment output produced by the fine-tuned Mistral model [51]. While the baseline model identified salient symptom patterns, its outputs were often verbose and interpretive. Fine-tuning resulted in more precise, consistently structured assessment representations, reducing ambiguity and improving alignment with the reference framework.

Conversation	Patient: I suddenly feel like I can't breathe and my heart races. Doctor: Do these happen without warning? Patient: Yes, and I worry about having another attack.
Real Observation	Panic Disorder (DSM-5 300.01)
Mistral Prediction (Before Fine-tuning)	<p>The symptoms mentioned in the conversation are:</p> <ol style="list-style-type: none"> 1. Sudden feelings of not being able to breathe (shortness of breath) 2. Rapid heart rate 3. Attacks occurring without warning 4. Patient worrying about having another attack <p>These symptoms align with the diagnostic criteria for Panic Disorder, as specified in the DSM-5:</p>
Mistral Prediction (After Fine-tuning)	Panic Disorder - DSM-5 300.01

Figure 12: Representative assessment output of the fine-tuned Mistral LLM aligned with structured reference labels (DSM-5 300.01).

Figure 13 shows a representative assessment output generated by the fine-

tuned Qwen2 model [52, 57]. Prior to fine-tuning, Qwen2 exhibited partial alignment with reference categories but occasionally lacked clarity and consistency. After fine-tuning, the model demonstrated improved mapping between observed symptom clusters and structured assessment labels, yielding outputs that were more concise, interpretable, and aligned with the reference framework.

Conversation	Patient: I hear voices that others can't hear. Doctor: Do you avoid certain places due to fear? Patient: Yes, I often avoid public spaces because I feel people are watching or trying to harm me.
Real Observation	Schizophrenia (DSM-5 295.90)
Qwen-2 Prediction (Before Fine-tuning)	<p>**Patient's symptoms:**</p> <ol style="list-style-type: none"> Hearing voices that others can't hear (auditory hallucinations) Avoiding certain places due to fear and perceived threat (avoidant behavior) <p>* Hearing voices that others can't hear is a classic symptom of auditory hallucinations, which is characteristic of Schizophrenia.</p> <p>* Avoiding public spaces due to perceived threat and fear suggests delusions (fixed, false beliefs) and paranoia, which are also common in Schizophrenia.</p>
Qwen-2 Prediction (After Fine-tuning)	Schizophrenia - DSM-5 295.90

Figure 13: Representative assessment output of the fine-tuned Qwen2 LLM aligned with structured reference labels (DSM-5 295.90).

Across all evaluated models, fine-tuning consistently improved predictive alignment, output structure, and interpretability. Similar qualitative improvements were observed across multiple assessment scenarios and reference conditions. Together, these results demonstrate that task-specific fine-tuning substantially enhances the suitability of LLMs as decision-support components within the Assessment Agent, supporting reliable and interpretable human-led assessment and escalation decisions under operational constraints [18].

5.5. Evaluation of Consensus-Driven Reasoning

This evaluation examines the performance of the reasoning-oriented LLM (GPT-OSS) in synthesizing and validating assessment outputs generated independently by multiple fine-tuned LLMs within the consortium [18]. The objective is to assess the system's ability to perform consensus-driven reasoning, reconciling diverse model perspectives into a coherent, interpretable, and reliable assessment representation suitable for human decision support.

In the evaluated workflow, individual fine-tuned LLMs generate structured assessment outputs based on the same input context, reflecting different reasoning emphases and interpretations. The reasoning-oriented LLM then evaluates these outputs, identifies areas of agreement and disagreement, and synthesizes a consolidated assessment accompanied by an explicit rationale. Figure 14 presents an illustrative example comparing the independent assessment outputs from the Llama-3, Mistral, and Qwen2 models with the final synthesized output produced by the reasoning LLM [14].

The results demonstrate that the reasoning-oriented LLM effectively resolves inconsistencies across model outputs by prioritizing coherence, contextual relevance, and alignment with the reference assessment framework. Rather than simply selecting a majority outcome, the reasoning model evaluates the supporting evidence provided by each model and produces a consolidated representation that reflects the most consistent and contextually grounded interpretation [58]. This process reduces single-model bias and mitigates the impact of spurious or overconfident predictions.

From a responsible and explainable AI perspective, the consensus-driven reasoning layer provides several key benefits. First, it introduces transparency by exposing intermediate model outputs and the rationale underlying the final synthesized assessment. Second, it improves robustness by ensuring that no single model’s prediction dominates the outcome. Third, it preserves human oversight by presenting the peer trainer with an interpretable assessment summary rather than an opaque or authoritative decision.

Importantly, the reasoning LLM operates strictly in an advisory capacity and does not perform autonomous diagnosis or intervention. Its role is limited to synthesizing assessment-level information to support human judgment within the Train-the-Trainers framework. The evaluation results indicate that incorporating a dedicated reasoning-oriented LLM as a consensus layer significantly enhances the reliability, interpretability, and trustworthiness of agentic AI decision support in sensitive and high-stakes environments [59].

Conversation	Patient: I keep getting flashbacks from the accident. Doctor: Do you avoid reminders of the event? Patient: Yes, I avoid driving and loud noises. (Sample 19)
Real Observation	Post-Traumatic Stress Disorder (DSM-5 309.81)
Prediction of Fine-tuned Llama-3 LLM	Post-Traumatic Stress Disorder
Prediction of Fine-tuned Mistral LLM	Post-Traumatic Stress Disorder (DSM-5 309.8)
Prediction of Fine-tuned Qwen-2 LLM	Unknown
Reasoning of OpenAI-o3 LLM	<p>Okay, let's tackle this mental disorder diagnosis problem. The user provided a doctor-patient conversation and predictions from three different LLMs. The task is to come up with a final diagnosis based on these predictions.</p> <p>First, I need to understand the conversation. The patient mentions having flashbacks from an accident and avoiding driving and loud noises. The doctor is asking about avoidance behaviors. The key symptoms here are flashbacks (intrusive memories) and avoidance of reminders. These are classic signs of PTSD.</p> <p>Looking at the model predictions: Llama-3 and Mistral both say PTSD, with Mistral even adding the DSM-5 code. Qwen says "Unknown." Why would Qwen say that? Maybe the model is more cautious or requires more symptoms to make a diagnosis. Let me recall the DSM-5 criteria for PTSD. The main criteria are exposure to trauma, intrusive symptoms (like flashbacks), avoidance, negative alterations in cognition/mood, and arousal/reactivity symptoms. The patient here has exposure (accident), intrusive memories, and avoidance. So they meet at least three criteria. But maybe Qwen is looking for more symptoms, like hypervigilance or sleep issues, which aren't mentioned here.</p> <p>But the two other models are confident. Llama-3 and Mistral's agreement strengthens the PTSD diagnosis. Qwen's uncertainty might be due to the absence of other symptoms in the conversation. However, according to DSM-5, you don't need all symptoms, just a subset from each category. The patient here has the necessary symptoms for a diagnosis.</p> <p>I should also check if there's any other possible disorder. For example, acute stress disorder is similar but occurs within a month of the trauma. Since the conversation doesn't specify the timeframe, but PTSD typically lasts longer. Also, the patient's avoidance and flashbacks are more aligned with PTSD.</p> <p>The final diagnosis should be PTSD, given the majority agreement and symptom alignment. Qwen's uncertainty might suggest the need for more information, but based on what's present, PTSD fits best.</p>

Figure 14: Consensus-driven assessment synthesis produced by the reasoning-oriented LLM through multi-model evaluation and reconciliation.

6. Related Work

Recent research has demonstrated the growing potential of large language models (LLMs) and AI-driven decision-support systems across a range of healthcare domains, including clinical reasoning, medical question answering, and diagnostic assistance. Systems such as Med-PaLM [26], MeLLaMA [60], and LLM-based differential diagnosis frameworks [28] have shown that domain-adapted language models can achieve strong performance on medical benchmarks and assist clinicians in structured reasoning tasks.

In mental health specifically, prior work has explored AI-assisted psychiatric assessment and diagnostic support using fine-tuned LLMs and hybrid reasoning pipelines [32]. Other systems, such as DrHouse [29], extend LLM-based decision support by incorporating multimodal inputs and iterative reasoning for primary care triage. However, most existing approaches are designed for civilian clinical settings, rely on centralized infrastructure, and assume direct clinician interaction. They typically emphasize autonomous or semi-autonomous diagnostic output rather than supporting peer-led intervention under operational constraints.

In parallel, advances in agentic AI and multi-agent systems have introduced new paradigms for automating complex workflows through coordinated reasoning, tool use, and iterative decision-making [23]. Recent studies have shown that agentic systems can leverage ensembles or consortia of LLMs combined with dedicated reasoning models to improve robustness, reduce single-model bias, and enhance interpretability [17, 38]. Consensus-driven reasoning approaches, including multi-model synthesis and structured deliberation, have proven particularly effective in high-uncertainty tasks [28]. Despite these advances, most agentic AI systems prioritize autonomy and task completion, with limited emphasis on strict human-in-the-loop control, explainability, or deployment in high-stakes and resource-constrained environments such as military operations.

In contrast, the proposed Train-the-Trainers framework uniquely integrates peer-based mental health support with a human-supervised, agentic AI architecture tailored for forward-deployed and austere settings. Rather than automating diagnosis or care delivery, the system augments recovered soldiers acting as peer trainers with structured decision support across assessment, peer intervention guidance, operational reasoning, and escalation. The use of a consensus-driven LLM consortium and a reasoning-oriented model enables transparent and explainable recommendations while preserving human authority and ethical boundaries. To the best of our knowledge, this work represents the first application of agentic AI to peer-led mental health support in battlefield environments, bridging gaps between military mental health practice, responsible AI design, and deployable multi-agent systems.

Table 5 summarizes key differences between prior AI-based healthcare and mental health systems and the proposed Train-the-Trainers framework.

Table 5: Comparison of AI-Based Healthcare and Mental Health Support Frameworks

Framework	Domain	Agentic AI	LLM Fine-Tuning	Supported LLMs	Human-in-the-Loop	Responsible AI	Explainable AI
Train-the-Trainers (This Work)	Battlefield Mental Health	✓	✓	LLaMA-3, Mistral, Qwen OpenAI-gpt-oss	✓	✓	✓
Med-PaLM [26]	General Medicine	✗	✓	PaLM	✓	✓	✗
LLM for DDx [28]	General Medicine	✗	✓	Multiple (unspecified)	✓	✗	✓
Me-LLaMA [60]	Biomedical NLP	✗	✓	LLaMA	✓	✗	✗
DrHouse [29]	Primary Care	✓	✓	Multiple (unspecified)	✓	✓	✓
CDSS [32]	Mental Health	✗	✗	GPT-4	✓	✓	✓
Weda-GPT [38]	Indigenous Medicine	✗	✓	LLaMA-3	✓	✓	✓

7. Conclusions and Future Work

This paper presented a Train-the-Trainers framework augmented by an agentic AI platform to support peer-based mental health assessment and intervention in forward-deployed and resource-constrained environments. By positioning recovered soldiers as peer trainers and equipping them with a human-supervised, multi-agent AI system, the proposed approach addresses critical gaps in timely psychological support where access to professional clinicians is limited. The framework emphasizes early assessment, structured peer support, informed escalation, and continuity of care while preserving clear boundaries between decision support and clinical authority.

We described the design and implementation of a functional prototype that integrates multiple specialized AI agents under a strict human-in-the-loop orchestration model. Central to the system is an Assessment Agent implemented using a consortium of fine-tuned large language models and a reasoning-oriented LLM that performs consensus-driven synthesis. The evaluation focused on the technical performance of this assessment pipeline, demonstrating that task-specific fine-tuning improves the consistency and structure of assessment outputs, while the consensus-driven reasoning layer enhances robustness, interpretability, and reliability. Importantly, all AI components operate in an advisory capacity, supporting human judgment rather than performing autonomous diagnosis or intervention.

The results indicate that agentic AI, when carefully designed with human oversight, explainability, and operational constraints in mind, can function as a force multiplier for peer-based mental health support in austere environments. The proposed framework is compatible with air-gapped deployment, emphasizes transparency and accountability, and aligns with responsible AI principles required for sensitive military and humanitarian applications.

Several directions for future work emerge from this study. First, broader evaluation using field-generated data and longitudinal studies is needed to assess system performance under real operational conditions and evolving

stress patterns. Second, future work will explore the integration of additional contextual signals, such as physiological or environmental indicators, while maintaining privacy and ethical safeguards. Third, user-centered evaluations involving peer trainers can provide insights into usability, trust, and cognitive load during high-stress interactions. Finally, extending the framework to support coordination with remote clinicians, secure telehealth interfaces, and cross-unit learning mechanisms represents a promising pathway for scaling the Train-the-Trainers approach across diverse operational contexts.

Overall, this work demonstrates the feasibility and potential of combining peer-based intervention models with consensus-driven, human-supervised agentic AI. By prioritizing responsible design, explainability, and human control, the proposed system offers a practical foundation for enhancing mental health support in environments where timely care is most difficult to deliver.

References

- [1] A. Flood and R. J. Keegan, “Cognitive resilience to psychological stress in military personnel,” *Frontiers in psychology*, vol. 13, p. 809003, 2022.
- [2] M. Deahl, M. Srinivasan, N. Jones, J. Thomas, C. Neblett, and A. Jolly, “Preventing psychological trauma in soldiers: The role of operational stress training and psychological debriefing,” *British Journal of Medical Psychology*, vol. 73, no. 1, pp. 77–85, 2000.
- [3] R. Gore, E. Bandara, S. Shetty, A. E. Musto, P. Rana, A. Valencia-Romero, C. Rhea, L. Tayebi, H. Richter, A. Yarlagadda *et al.*, “Proof-of-tbi–fine-tuned vision language model consortium and openai-o3 reasoning llm-based medical diagnosis support system for mild traumatic brain injury (tbi) prediction,” *arXiv preprint arXiv:2504.18671*, 2025.
- [4] J. M. Orasanu and P. Backer, “Stress and military performance,” in *Stress and human performance*. Psychology Press, 2013, pp. 89–125.
- [5] V. Svetlitzky, M. Farchi, A. B. Yehuda, and A. B. Adler, “Witnessing acute stress reaction in team members: The moderating effect of peer-based training,” *The Journal of Nervous and Mental Disease*, vol. 208, no. 10, pp. 803–809, 2020.

- [6] A. B. Adler, I. Gutierrez, H. M. Edge, A. Nordstrand, A. Simms, and G. Willmund, “Peer-based intervention for acute stress reaction: adaptations by five militaries,” *BMJ Mil Health*, vol. 170, no. 5, pp. 425–429, 2024.
- [7] D. B. Acharya, K. Kuppan, and B. Divya, “Agentic ai: Autonomous intelligence for complex goals—a comprehensive survey,” *IEEE Access*, 2025.
- [8] M. Gridach, J. Nanavati, K. Z. E. Abidine, L. Mendes, and C. Mack, “Agentic ai for scientific discovery: A survey of progress, challenges, and future directions,” *arXiv preprint arXiv:2503.08979*, 2025.
- [9] E. Bandara, R. Gore, P. Foytik, S. Shetty, R. Mukkamala, A. Rahman, X. Liang, S. H. Bouk, A. Hass, S. Rajapakse *et al.*, “A practical guide for designing, developing, and deploying production-grade agentic ai workflows,” *arXiv preprint arXiv:2512.08769*, 2025.
- [10] E. Bandaraa, T. Hewab, S. Rajapaksef, I. Kularathnaf, P. Karunarathnag, R. Gorea, P. Foytika, S. Shettya, R. Mukkamalaa, A. Rahmanc *et al.*, “Rovanima—scaling up small and medium-sized tourism enterprises through agentic ai in lapland.”
- [11] J. Alshaer and I. R. Amman, “Mitigating data leakage in high-compliance environments: A governance framework for local-llm adoption,” 2026.
- [12] I. H. Sarker, “Llm potentiality and awareness: a position paper from the perspective of trustworthy and responsible ai modeling,” *Discover Artificial Intelligence*, vol. 4, no. 1, p. 40, 2024.
- [13] R. Dwivedi, D. Dave, H. Naik, S. Singhal, R. Omer, P. Patel, B. Qian, Z. Wen, T. Shah, G. Morgan *et al.*, “Explainable ai (xai): Core ideas, techniques, and solutions,” *ACM computing surveys*, vol. 55, no. 9, pp. 1–33, 2023.
- [14] E. Bandara, T. Hewa, R. Gore, S. Shetty, R. Mukkamala, P. Foytik, A. Rahman, S. H. Bouk, X. Liang, A. Hass *et al.*, “Towards responsible and explainable ai agents with consensus-driven reasoning,” *arXiv preprint arXiv:2512.21699*, 2025.

- [15] S. Arora, B. Yang, S. Eyuboglu, A. Narayan, A. Hojel, I. Trummer, and C. Ré, “Language models enable simple systems for generating structured views of heterogeneous data lakes,” *arXiv preprint arXiv:2304.09433*, 2023.
- [16] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, “Language models are few-shot learners,” *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.
- [17] Y. Zhang, S. Mao, T. Ge, X. Wang, A. de Wynter, Y. Xia, W. Wu, T. Song, M. Lan, and F. Wei, “Llm as a mastermind: A survey of strategic reasoning with large language models,” *arXiv preprint arXiv:2404.01230*, 2024.
- [18] J. Wang, “A tutorial on llm reasoning: Relevant methods behind chatgpt o1,” *arXiv preprint arXiv:2502.10867*, 2025.
- [19] X. Lin, W. Wang, Y. Li, S. Yang, F. Feng, Y. Wei, and T.-S. Chua, “Data-efficient fine-tuning for llm-based recommendation,” in *Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval*, 2024, pp. 365–374.
- [20] I. Shruti, A. Kumar, A. Seth *et al.*, “Responsible generative ai: A comprehensive study to explain llms,” in *2024 International Conference on Electrical, Computer and Energy Technologies (ICECET)*. IEEE, 2024, pp. 1–6.
- [21] I. Arous, K. Chehbouni, Z. Cheng, and B. Dossou, “Llm explainability,” in *Handbook of Human-Centered Artificial Intelligence*. Springer, 2025, pp. 1–61.
- [22] A. Yehudai, L. Eden, A. Li, G. Uziel, Y. Zhao, R. Bar-Haim, A. Cohan, and M. Shmueli-Scheuer, “Survey on evaluation of llm-based agents,” *arXiv preprint arXiv:2503.16416*, 2025.
- [23] E. Bandara, R. Gore, X. Liang, S. Rajapakse, I. Kularathne, P. Karunarathna, P. Foytik, S. Shetty, R. Mukkamala, A. Rahman *et al.*, “Agentsway—software development methodology for ai agents-based teams,” *arXiv preprint arXiv:2510.23664*, 2025.

- [24] Y. Auxéméry, “Post-traumatic psychiatric disorders: Ptsd is not the only diagnosis,” *La Presse Médicale*, vol. 47, no. 5, pp. 423–430, 2018.
- [25] E. Bandara, R. Gore, A. Yarlagadda, A. H. Clayton, P. Samuel, C. K. Rhea, and S. Shetty, “Standardization of psychiatric diagnoses—role of fine-tuned llm consortium and openai-gpt-oss reasoning llm enabled decision support system,” *arXiv preprint arXiv:2510.25588*, 2025.
- [26] K. Singhal, S. Azizi, T. Tu, S. S. Mahdavi, J. Wei, H. W. Chung, N. Scales, A. Tanwani, H. Cole-Lewis, S. Pfohl *et al.*, “Large language models encode clinical knowledge,” *Nature*, vol. 620, no. 7972, pp. 172–180, 2023.
- [27] S. Bračko and A. Čelofiga, ““it’s a psychiatric patient”: Misdiagnosing of somatic symptoms in patients with mental disorders due to stigma and inadequate diagnostic treatment,” *Archives of Psychiatry Research: An International Journal of Psychiatry and Related Sciences*, vol. 60, no. 1., pp. 62–66, 2024.
- [28] D. McDuff, M. Schaekermann, T. Tu, A. Palepu, A. Wang, J. Garrison, K. Singhal, Y. Sharma, S. Azizi, K. Kulkarni *et al.*, “Towards accurate differential diagnosis with large language models,” *arXiv preprint arXiv:2312.00164*, 2023.
- [29] B. Yang, S. Jiang, L. Xu, K. Liu, H. Li, G. Xing, H. Chen, X. Jiang, and Z. Yan, “Drhouse: An llm-empowered diagnostic reasoning system through harnessing outcomes from sensor data and expert knowledge,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 4, pp. 1–29, 2024.
- [30] K. D. Kannan, S. K. Jagatheesaperumal, R. N. Kandala, M. Lotfaliany, R. Alizadehsanid, and M. Mohebbi, “Advancements in machine learning and deep learning for early detection and management of mental health disorder,” *arXiv preprint arXiv:2412.06147*, 2024.
- [31] C. Yin, F. Li, S. Zhang, Z. Wang, J. Shao, P. Li, J. Chen, and X. Jiang, “Mdd-5k: A new diagnostic conversation dataset for mental disorders synthesized via neuro-symbolic llm agents,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 24, 2025, pp. 25 715–25 723.

- [32] B. H. Kim and C. Wang, “Large language models for interpretable mental health diagnosis,” *arXiv preprint arXiv:2501.07653*, 2025.
- [33] H. Samo, K. Ali, M. Memon, F. A. Abbasi, M. Y. Koondhar, and K. Dahri, “Fine-tuning mistral 7b large language model for python query response and code generation: A parameter efficient approach,” *VAWKUM Transactions on Computer Sciences*, vol. 12, no. 1, pp. 205–217, 2024.
- [34] E. Bandara, S. Shetty, R. Mukkamala, A. Rahman, P. Foytik, X. Liang, K. De Zoysa, and N. W. Keong, “Devsec-gpt — generative-ai (with custom-trained meta’s llama2 llm), blockchain, nft and pbom enabled cloud native container vulnerability management and pipeline verification platform,” in *2024 IEEE Cloud Summit*, 2024, pp. 28–35.
- [35] A. Bandi, B. Kongari, R. Naguru, S. Pasnoor, and S. V. Vilipala, “The rise of agentic ai: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges,” *Future Internet*, vol. 17, no. 9, p. 404, 2025.
- [36] E. Bandara, S. H. Bouk, S. Shetty, S. Roy, R. Mukkamala, A. Rahman, P. Foytik, X. Liang, N. W. Keong, and K. De Zoysa, “Llama-recipe — fine-tuned meta’s llama llm, pbom and nft enabled 5g network-slice orchestration and end-to-end supply-chain verification platform,” in *2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC)*, 2025, pp. 1–6.
- [37] A. Augustin, J. Yi, T. Clausen, and W. Townsley, “A study of lora: Long range & low power networks for the internet of things,” *Sensors*, vol. 16, no. 9, p. 1466, 2016.
- [38] E. Bandara, P. Foytik, S. Shetty, R. Mukkamala, A. Rahman, X. Liang, N. W. Keong, and K. De Zoysa, “Wedagpt—generative-ai (with custom-trained meta’s llama2 llm), blockchain, self sovereign identity, nft and model card enabled indigenous medicine platform,” in *2024 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2024, pp. 1–6.
- [39] T. Raheem and G. Hossain, “Agentic ai systems: Opportunities, challenges, and trustworthiness,” in *2025 IEEE International Conference on Electro Information Technology (eIT)*. IEEE, 2025, pp. 618–624.

- [40] E. Bandara, A. Hass, S. Shetty, R. Mukkamala, R. Gore, A. Rahman, and S. H. Bouk, “Deep-stride: Automated security threat modeling with vision-language models,” in *2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2025, pp. 1–7.
- [41] M. M. Hasan, H. Li, E. Fallahzadeh, G. K. Rajbahadur, B. Adams, and A. E. Hassan, “Model context protocol (mcp) at first glance: Studying the security and maintainability of mcp servers,” *arXiv preprint arXiv:2506.13538*, 2025.
- [42] X. Hou, Y. Zhao, S. Wang, and H. Wang, “Model context protocol (mcp): Landscape, security threats, and future research directions,” *arXiv preprint arXiv:2503.23278*, 2025.
- [43] E. Bandara, S. Shetty, R. Mukkamala, R. Gore, P. Foytik, S. H. Bouk, A. Rahman, X. Liang, N. W. Keong, K. De Zoysa *et al.*, “Model context contracts-mcp-enabled framework to integrate llms with blockchain smart contracts,” *arXiv preprint arXiv:2510.19856*, 2025.
- [44] Y. Lee, S. Lee, J. Choi, and W. Jung, “Implementation of an on-device ai chatbot system in an air-gapped control central console environment,” *Journal of the Korea Institute of Military Science and Technology*, vol. 28, no. 5, pp. 525–535, 2025.
- [45] Y. Zheng, R. Zhang, J. Zhang, Y. Ye, Z. Luo, Z. Feng, and Y. Ma, “Llamafactory: Unified efficient fine-tuning of 100+ language models,” *arXiv preprint arXiv:2403.13372*, 2024.
- [46] C. Liao, M. Sun, Z. Yang, J. Xie, K. Chen, B. Yuan, F. Wu, and Z. Wang, “Lohan: Low-cost high-performance framework to fine-tune 100b model on a consumer gpu,” *arXiv preprint arXiv:2403.06504*, 2024.
- [47] H. Kimm, I. Paik, and H. Kimm, “Performance comparison of tpu, gpu, cpu on google colaboratory over distributed deep learning,” in *2021 IEEE 14th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoc)*. IEEE, 2021, pp. 312–319.
- [48] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, “Qlora: Efficient finetuning of quantized llms,” *Advances in Neural Information Processing Systems*, vol. 36, 2024.

- [49] E. Bandara, S. H. Bouk, S. Shetty, R. Gore, S. Kompella, R. Mukkamala, A. Rahman, P. Foytik, X. Liang, N. W. Keong, and K. De Zoysa, “Vindsec-llama — fine-tuned meta’s llama-3 llm, federated learning, blockchain and pbom-enabled data security architecture for wind energy data platforms,” in *2025 International Wireless Communications and Mobile Computing (IWCMC)*, 2025, pp. 120–126.
- [50] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan *et al.*, “The llama 3 herd of models,” *arXiv preprint arXiv:2407.21783*, 2024.
- [51] P. Agrawal, S. Antoniak, E. B. Hanna, B. Bout, D. Chaplot, J. Chudnovsky, D. Costa, B. De Monicault, S. Garg, T. Gervet *et al.*, “Pixtral 12b,” *arXiv preprint arXiv:2410.07073*, 2024.
- [52] P. Wang, S. Bai, S. Tan, S. Wang, Z. Fan, J. Bai, K. Chen, X. Liu, J. Wang, W. Ge *et al.*, “Qwen2-vl: Enhancing vision-language model’s perception of the world at any resolution,” *arXiv preprint arXiv:2409.12191*, 2024.
- [53] S. Agarwal, L. Ahmad, J. Ai, S. Altman, A. Applebaum, E. Arbus, R. K. Arora, Y. Bai, B. Baker, H. Bao *et al.*, “gpt-oss-120b & gpt-oss-20b model card,” *arXiv preprint arXiv:2508.10925*, 2025.
- [54] D. A. Regier, E. A. Kuhl, and D. J. Kupfer, “The dsm-5: Classification and criteria changes,” *World psychiatry*, vol. 12, no. 2, pp. 92–98, 2013.
- [55] M. S. Scheeringa, “Is factor analysis useful for revising diagnostic criteria for ptsd? a systematic review of five issues ten years after dsm-5,” *Journal of Psychiatric Research*, 2024.
- [56] R. E. Sistrup, R. Kimerling, P. P. Schnurr, and M. J. Bovin, “The impact of screening positive for hazardous alcohol use on the diagnostic accuracy of the ptsd checklist for dsm-5 among veterans,” *Journal of Traumatic Stress*, vol. 37, no. 2, pp. 328–336, 2024.
- [57] M. Xiang, R. Fernando, and B. Wang, “On-device qwen2. 5: Efficient llm inference with model compression and hardware acceleration,” *arXiv preprint arXiv:2504.17376*, 2025.

- [58] E. Bandara, R. Gore, S. Shetty, R. Mukkamala, C. Rhea, A. Yarlagadda, S. Kaushik, L. De Silva, A. Maznychenko, I. Sokolowska *et al.*, “Standardization of neuromuscular reflex analysis—role of fine-tuned vision-language model consortium and openai gpt-oss reasoning llm enabled decision support system,” *arXiv preprint arXiv:2508.12473*, 2025.
- [59] A. Singh, A. Ehtesham, S. Kumar, and T. T. Khoei, “Enhancing ai systems with agentic workflows patterns in large language model,” in *2024 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2024, pp. 527–532.
- [60] Q. Xie, Q. Chen, A. Chen, C. Peng, Y. Hu, F. Lin, X. Peng, J. Huang, J. Zhang, V. Keloth *et al.*, “Me-llama: Medical foundation large language models for comprehensive text analysis and beyond,” 2024.