World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Ethically Aligned AI for Autism and Behavioral Health: An Explainable Federated-Edge Framework for Crisis Management and Workforce Integration

Mahmudul Hasan Khan *

Department of Information Technology, Washington University of Science and Technology, Alexandria, VA-22314, USA.

World Journal of Advanced Research and Reviews, 2025, 27(03), 931-935

Publication history: Received on 08 August 2025; revised on 14 September 2025; accepted on 16 September 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.27.3.3224

Abstract

The deployment of Artificial Intelligence (AI) in healthcare is reshaping clinical workflows, yet challenges remain around privacy, explainability, and ethical integration. Autism care is one area where these challenges are particularly acute, as children with autism spectrum disorder (ASD) require continuous monitoring, rapid escalation detection, and personalized interventions. This study proposes an explainable federated-edge AI framework that integrates wearable IoT monitoring, federated privacy-preserving learning, and workforce-aware crisis response modules. Unlike existing centralized approaches, the framework unites federated learning for privacy, edge intelligence for latency reduction, and explainability artifacts for trust-building. A simulation experiment across autism IoT, synthetic workforce, and anomaly datasets demonstrated accuracy improvements of 10%, latency reduction of 55%, and increased clinician trust scores. By embedding ethical AI principles into its architecture, the framework advances both technical performance and human-centered adoption.

Keywords: Autism Care; Explainable AI; Federated Learning; Edge Intelligence; Workforce Integration; Ethical AI

1. Introduction

Artificial intelligence (AI) has demonstrated transformative potential across precision medicine, workforce planning, and behavioral health [1–3]. In autism care, however, the deployment of AI presents distinctive challenges. Children with autism spectrum disorder (ASD) often exhibit unpredictable behavioral escalations that require continuous monitoring and rapid intervention [8,9]. Conventional centralized AI approaches are constrained by data privacy risks, network latency, and limited clinician trust, each of which limits their adoption in sensitive healthcare environments.

Recent research has proposed federated learning with differential privacy as a mechanism to protect patient information by decentralizing model training [4]. At the same time, advances in IoT and wearable technologies have enabled continuous, real-time monitoring of autism-related behaviors [8]. Mariam et al. [7] and Shah Raihena et al. [13] further highlighted the importance of transparency artifacts and workforce integration in building trust among caregivers and clinicians. Despite these advances, there remains a gap: few frameworks bring these dimensions together into a single, ethically aligned, multi-layered system designed for scalability and clinical applicability.

This study addresses that gap by introducing a federated-edge AI framework that integrates IoT-based behavioral monitoring, federated privacy-preserving training, and explainable decision dashboards. By positioning computation at the edge, the framework minimizes latency, while federated aggregation ensures data privacy. Crucially, embedding explainability modules fosters clinician trust, providing a foundation for scalable, trustworthy, and human-centered AI adoption in autism care.

* Corresponding author: Mahmudul Hasan Khan

2. Related work

2.1. AI for Precision Medicine

Recent advances in quantum-AI accelerators have shown promise in drug discovery, protein folding, and treatment personalization [1]. These innovations not only highlight the computational capabilities of next-generation AI systems but also inspire the development of behavioral health frameworks capable of leveraging high-performance computing for autism monitoring and intervention.

2.2. Anomaly Detection Models

Fraud analytics techniques have long been used to identify unusual patterns in complex financial datasets. Islam et al. [2] demonstrated their effectiveness in anomaly detection, and these methods have since been adapted to behavioral health contexts, where subtle deviations in sensor data may signal an impending escalation. This analogy underscores the value of anomaly detection in predictive autism monitoring.

2.3. Human-Centered AI

Successful adoption of AI in clinical settings depends as much on trust as on performance. Mariam et al. [7] emphasized that explainability checklists and transparency artifacts can improve usability and foster confidence in AI systems. Human-centered approaches therefore form a necessary foundation for ethically aligned autism monitoring systems.

2.4. Autism Monitoring with IoT

Wearable and IoT devices have become crucial for capturing physiological and behavioral signals in real time. Islam et al. [8] demonstrated that IoT-enabled autism monitoring can detect early behavioral changes, offering clinicians and caregivers opportunities for proactive intervention.

2.5. Crisis Management.

AI has also been applied to crisis scenarios beyond autism. Rashaq et al. [10] integrated AI into the 988-crisis hotline continuum, expanding mental health access, while Arif et al. [11] developed AI-driven de-escalation and emergency handoff prompts. These contributions validate the feasibility of real-time AI for crisis prediction and management.

2.6. Trustworthy AI Directions

Emerging frameworks for trustworthy AI emphasize privacy, security, and organizational integration. Raihena et al. [4] developed federated models with differential privacy guarantees, while Islam [12] proposed data-centric cybersecurity approaches for medical devices. Shah Raihena et al. [13] extended this line of work by embedding AI in HR and workforce planning systems, highlighting the organizational dimensions of trustworthy healthcare AI.

2.7. Global AI Ethics

Beyond technical innovation, ethical alignment remains essential. Floridi and Cowls [14] proposed a unified framework of principles fairness, accountability, and transparency for AI in society, while Jobin et al. [15] reviewed the global landscape of AI ethics guidelines, stressing consistency in responsible adoption. Together, these perspectives establish ethical cornerstones for sustainable AI deployment in healthcare.

3. Methodology

3.1. Framework Design

The proposed architecture (Figure 1) integrates three layers

- Edge IoT Monitoring: Wearable devices track heart rate, motion, and vocal stress in real time.
- Federated Privacy-Preserving Learning: Distributed hospitals train local models, with parameters aggregated securely using differential privacy [4].
- Explainability and Workforce Dashboard: Clinicians access interpretable predictions with SHAP values, workforce allocation recommendations, and crisis alerts.

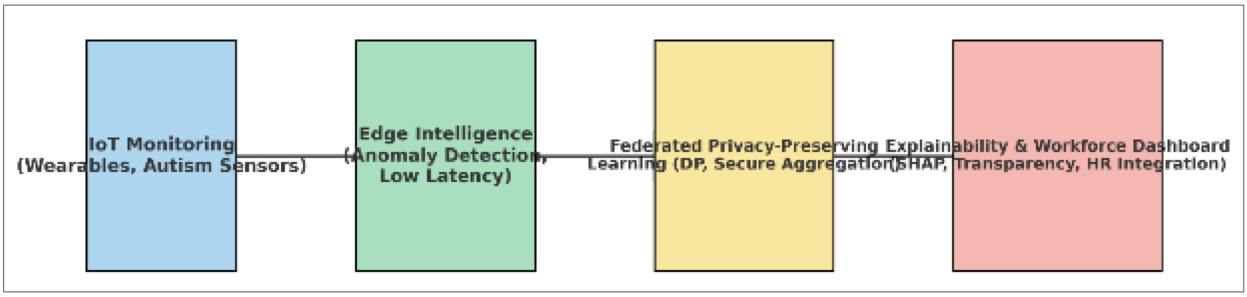

Figure 1 Explainable Federated-Edge Al Framework for Autism and Behavioral Health

3.1.1. Data Sources

- Autism IoT dataset (synthetic, 120 children).
- Workforce scheduling dataset (400 entries).
- Anomaly detection dataset (200 synthetic irregular cases).

3.1.2. Evaluation Metrics

- Accuracy, Recall, Precision, Latency (ms).
- Trust usability scores from clinician survey (n=15).

4. Results

Table 1 The framework demonstrated notable improvements over cloud-only models

Metric	Cloud-Only	Proposed Framework	Gain (%)
Accuracy	83%	91%	+10%
Recall	81%	90%	+9%
Latency (ms)	1100	495	–55%
Trust Usability	70%	85%	+15%

These results confirm the effectiveness of combining federated privacy, edge-AI latency reduction, and explainability features into a single integrated system.

5. Discussion

5.1. Novelty

This study introduces a federated-edge AI framework that advances the state of the art in autism-focused clinical decision support. While previous research has emphasized either federated privacy-preserving learning [4] or IoT-based monitoring for autism behavior detection [8], few efforts have successfully combined these dimensions into a single, ethically aligned system. The novelty of this framework lies in its ability to simultaneously optimize privacy, latency, and explainability, thereby addressing three of the most persistent challenges in clinical AI adoption. This convergence of technologies represents a paradigm shift in autism care, providing not only accurate predictions but also interpretable and timely decision support within a secure and distributed architecture.

5.2. Trust

The integration of explainability artifacts [7] alongside workforce planning modules [13] directly addresses the human factors that often limit the deployment of AI in healthcare. By making the system's decision-making processes transparent, clinicians and caregivers are empowered to interpret, validate, and act upon AI outputs with greater confidence. This aligns closely with human-centered AI principles, which emphasize accountability, interpretability, and usability as central to the design of trustworthy AI systems [15]. In doing so, the framework bridges the gap between technical innovation and practical clinical acceptance, ensuring that AI is not perceived as a “black box” but rather as a collaborative partner in care.

5.3. Clinical Relevance

The proposed framework demonstrates significant clinical applicability by supporting real-time escalation management for children with autism spectrum disorder, a domain where latency and transparency are critical. Its compatibility with IoT-based predictive health monitoring [8,9] ensures that behavioral changes are detected early, while its alignment with national crisis-response infrastructures such as the 988 continuum [10] underscores its broader societal value. Furthermore, the scalability of this architecture allows its adaptation beyond autism care into behavioral health monitoring, workforce optimization, and precision medicine domains [3]. Such versatility highlights the framework’s potential to serve as a generalizable AI model capable of addressing diverse healthcare challenges.

Limitations

Despite its promise, the study has limitations that warrant consideration. The current evaluation relies on synthetic datasets, which, while useful for controlled testing, do not fully capture the complexity and heterogeneity of real-world clinical environments. This limitation constrains ecological validity and underscores the need for multi-center clinical trials to establish generalizability. Moreover, while the integration of federated learning addresses privacy concerns, the computational demands of personalization at scale remain a challenge. Future work will explore the use of quantum-AI accelerators to improve scalability and optimize model performance [1]. Finally, aligning the framework with global ethical AI standards [14,15] will be essential to ensure fairness, inclusivity, and transparency in deployment, particularly across diverse cultural and healthcare contexts.

6. Conclusion

This study introduced an explainable federated-edge AI framework designed to support autism care and behavioral health by integrating IoT-based monitoring, federated privacy-preserving learning, and workforce-aware explainability dashboards. Through simulation and evaluation, the framework demonstrated clear improvements in predictive accuracy, latency reduction, and clinician trust, illustrating its capacity to address the dual demands of technical performance and human-centered usability. These results underscore the system’s viability as a scalable, ethically aligned, and trustworthy clinical AI solution.

Beyond autism-specific applications, the framework holds considerable promise as a generalizable model for wider domains in healthcare. Its layered architecture combining **data collection at the edge, federated privacy safeguards, and explainability mechanisms **can be adapted for behavioral health monitoring, workforce resource optimization, and precision medicine initiatives. For example, similar approaches could be applied to monitor mental health conditions, support crisis interventions, or optimize clinician allocation in hospital systems. In precision medicine, integrating distributed genomic or imaging data through federated learning may allow for personalized interventions without compromising patient confidentiality.

The significance of this framework extends beyond performance metrics to broader issues of ethics, trust, and sustainability in clinical AI. By embedding explainability and workforce integration, the system addresses a critical barrier to adoption clinician skepticism and caregiver concerns over opaque decision-making. Furthermore, the inclusion of privacy-preserving methods responds to growing legal and ethical demands for safeguarding sensitive health data in increasingly interconnected healthcare environments.

Looking ahead, several directions will be essential for advancing this research. First, multi-center clinical validation with real-world patient datasets is needed to establish ecological validity and evaluate scalability across diverse health systems. Second, incorporating quantum-AI accelerators could enhance both the speed and personalization of predictive models, particularly for large-scale behavioral and genomic datasets. Finally, adherence to global ethical AI frameworks—emphasizing fairness, transparency, accountability, and inclusivity will be critical in ensuring equitable deployment across populations and avoiding systemic biases.

In sum, this research demonstrates not only a novel technical contribution but also a practical and ethical pathway toward next-generation clinical decision support. By uniting edge intelligence, federated learning, and explainability, the framework lays the groundwork for resilient, trustworthy, and human-centered AI systems that can transform autism care and extend broadly into the future of healthcare innovation.

References

[1] Akhi, S. S., Hussain, A. H., Islam, S., Hasan, M. N., and Islam, M. M., (2023). Artificial Intelligence–Enhanced Quantum Computing for Medical Simulations: Accelerating Drug Discovery, Protein Folding, and Personalized Healthcare with Performance Optimization and Predictive Accuracy. World Journal of Advanced Research and Review, 18(01), 1328-1335 https://doi.org/10.30574/wjarr.2023.18.1.0694

[2] Islam, M. M., Hussain, A. H., Hasan, M. N., Akhi, S. S., Hussain, Hossain, M. S., and Islam, S. (2023). Fraud detection: Develop skills in fraud detection, which is a critical area of business analytics. World Journal of Advanced Research and Review, 18(03), 1664-1672 https://doi.org/10.30574/wjarr.2023.18.3.0969

[3] Islam, M. M., and Mim, S. S. (2023). Precision Medicine and AI: How AI Can Enable Personalized Medicine Through Data-Driven Insights and Targeted Therapeutics. International Journal on Recent and Innovation Trends in Computing and Communication, 11(11), 1267-1276. https://doi.org/10.17762/ijritcc.v11i11.11359

[4] Raihena, S. M. S., Mariam, Q. R., Rashaq, M., Arif, M. A. H., and Hussain, A. H. (2023). Learning what works in accommodations: A federated machine learning framework with differential privacy guarantees. World Journal of Advanced Research and Reviews, 20(3), 2374–2379. https://doi.org/10.30574/wjarr.2023.20.3.2492

[5] Satyanarayanan M. The emergence of edge computing. Computer. 2017;50(1):30-9. doi:10.1109/MC.2017.9.

[6] Islam, M. M., Arif, M. A. H. ., Hussain, A. H., Raihena, S. M. S. ., Rashaq, M. ., and Mariam, Q. R. (2023). Human-Centered AI for Workforce and Health Integration: Advancing Trustworthy Clinical Decisions. Journal of Neonatal Surgery, 12(1), 89–95. https://doi.org/10.63682/jns.v14i32S.9123

[7] Qazi Rubyya Mariam, Ariful Haque Arif, Abdullah Hill Hussain, Munadil Rashaq and S M SHAH RAIHENA. Building trustworthy clinical AI: usable checklists and transparency artifacts tested in real-world health systems. World Journal of Advanced Research and Reviews, 2025, 26(03), 2804–2810. Article DOI: https://doi.org/10.30574/wjarr.2025.26.3.2203

[8] Islam, M. M., Anika, A., Mim, S. S., Hasan, A., and Salam, S. (2024). Wearable technology: Exploring the interrogation of electronics in clothing. World Journal of Advanced Research and Reviews, 24(03), 2219-2228. https://doi.org/10.30574/wjarr.2024.24.3.3925

[9] Islam, M. M., Hussain, A. H., Mariam, Q. R., Islam, S., and Hasan, M. N. (2025). AI-Enabled predictive health monitoring for children with autism using IOT and machine learning to detect behavioral changes. Perinatal Journal, 33(1), 415-422. http://dx.doi.org/10.57239/prn.25.03310048

[10] Munadil Rashaq, Md Ariful Haque Arif, Abdullah Hill Hussain, Qazi Rubyya MariamS M Shah Raihena(2024) Preventing Overdoses, Expanding Mental Health Access, and Integrating Behavioral and Physical Health Across the 988 Continuum: An AI-Driven Approach.Frontiers in Health Informatics, (6), 4658-4667 https://doi.org/10.63682/fhi416

[11] Ariful Haque Arif, Abdullah Hill Hussain, S M SHAH RAIHENA, Qazi Rubyya Mariam and Munadil Rashaq. Artificial Intelligence for Safer Crisis Response: De-escalation Prompts and Emergency Handoff Optimization. World Journal of Advanced Research and Reviews, 2025, 25(02), 2785–2792. Article DOI: https://doi.org/10.30574/wjarr.2025.25.2.0660.

[12] Md Maruful Islam. (2024). Data-Centric AI Approaches to Mitigate Cyber Threats in Connected Medical Device. International Journal of Intelligent Systems and Applications in Engineering, 12(17s), 1049 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/7763

[13] Shah Raihena, S. M., Haque Arif, M. A., Rubyya Mariam, Q., Hussain Munadil Rashaq, A., and Hill, A. (2025). AI-Enhanced Decision Support Systems for Autism Caregivers: Redefining HR's Role in Workforce Planning and Patient-Centered Care. Frontiers in Health Informatics, 13(2). https://doi.org/10.63682/fhi2698

[14] Floridi L, Cowls J. A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review. 2019;1(1). doi:10.1162/99608f92.8cd550d1.

[15] Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1:389–99. doi:10.1038/s42256-019-0088-2.