Frontiers in Computer Science and Artificial Intelligence
ISSN: 2978-8048
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Operationalizing the NIST AI Risk Management Framework for AI-Driven Autism Care Systems: A Governance-Centric Technical Architecture

M Salman Khan
Department of Computer Science & Engineering, Brac University, Dhaka, Bangladesh
Corresponding Author: M Salman Khan**, E-mail**: salmankhan0006@gmail.com

| ABSTRACT

Artificial intelligence–driven systems are increasingly deployed in autism spectrum disorder care to support behavioral monitoring, escalation prediction, and caregiver decision-making. While these systems offer substantial clinical and operational benefits, they also introduce risks related to safety, privacy, bias, accountability, and trust—particularly when applied to vulnerable pediatric populations. The NIST Artificial Intelligence Risk Management Framework provides high-level guidance for managing AI risks, yet practical operationalization within real-world autism care systems remains limited. This study proposes a governance-centric technical architecture that embeds NIST AI Risk Management Framework principles directly into the design, deployment, and lifecycle management of AI-driven autism care systems. The framework translates abstract governance functions into concrete technical controls, audit artifacts, and operational workflows aligned with reinforcement learning, IoT-based behavioral monitoring, and caregiver decision support. Through a structured risk taxonomy, control catalog, and implementation blueprint, this research demonstrates how trustworthy AI principles can be transformed into deployable, auditable, and scalable autism care solutions.

| KEYWORDS

Autism spectrum disorder; AI governance; NIST AI Risk Management Framework; Trustworthy AI; Clinical decision support; AI safety; Risk-aware system architecture

| ARTICLE INFORMATION

ACCEPTED: 01 January 2026 **PUBLISHED:** 04 January 2026 **DOI:** 10.32996/jcsts.2026.5.1.3

Introduction

Artificial intelligence (AI) has rapidly transitioned from experimental research to operational deployment across healthcare and assistive technology domains. In autism spectrum disorder (ASD) care, AI-driven systems are increasingly used to support behavioral monitoring, escalation prediction, personalized intervention planning, and caregiver decision support. These systems leverage machine learning, reinforcement learning, and Internet of Things (IoT) infrastructures to process continuous streams of behavioral, physiological, and contextual data.

Despite their promise, AI-driven autism care systems operate in environments characterized by high sensitivity, ethical complexity, and significant human impact. Children with autism represent a vulnerable population, and AI-influenced decisions may directly affect emotional well-being, safety, and long-term developmental outcomes. Errors, bias, misinterpretation, or misuse of AI systems can lead to inappropriate interventions, increased distress, or erosion of caregiver trust.

Historically, AI development has prioritized performance metrics such as accuracy, latency, and scalability. While these metrics remain important, they are insufficient when AI systems interact directly with human caregivers and children. Increasingly, researchers and regulators recognize that AI systems must be **trustworthy by design**, incorporating governance, accountability, and risk management throughout their lifecycle.

The National Institute of Standards and Technology introduced the **Artificial Intelligence Risk Management Framework (AI RMF)** to provide a structured approach for identifying, assessing, and managing AI-related risks. The framework emphasizes four core functions—**Govern, Map, Measure, and Manage**—intended to guide organizations in deploying AI responsibly across diverse application domains.

However, the AI RMF is intentionally high-level and technology-agnostic. While this flexibility allows broad applicability, it also creates a significant implementation gap. Practitioners deploying AI in specialized domains such as autism care often struggle to translate abstract governance principles into concrete technical controls, system architectures, and operational workflows.

As a result, governance is frequently treated as an external compliance activity rather than an integral component of AI system design. Policies may exist on paper, but they are rarely embedded into model training pipelines, decision logic, caregiver interfaces, or monitoring processes. This disconnect undermines the effectiveness of governance frameworks and limits their impact on real-world AI behavior.

This study addresses this gap by proposing a **governance-centric technical architecture** that operationalizes the NIST AI Risk Management Framework within AI-driven autism care systems. Rather than layering governance on top of existing systems, the proposed approach embeds risk management directly into data ingestion, learning algorithms, decision support interfaces, and lifecycle monitoring.

The objectives of this research are:

1. To identify and categorize AI-specific risks in autism care systems.
2. To translate NIST AI RMF functions into concrete architectural components and workflows.
3. To propose a deployable and auditable governance-centric system architecture.
4. To demonstrate how governance integration enhances trust, safety, and sustainability without compromising system performance.

Background and Related Work

AI-Driven Autism Care Systems

AI applications in autism care span multiple functional areas, including behavioral escalation prediction, personalized monitoring, and caregiver decision support. Reinforcement learning has been shown to effectively model escalation as a temporal process, enabling early intervention based on learned state transitions **[1]**. IoT-based frameworks support continuous behavioral tracking across home and clinical environments **[2,4,9]**, while AI-augmented decision support systems translate predictions into caregiver-facing recommendations **[5,10]**.

While technically effective, these systems often prioritize predictive performance and operational efficiency. Governance considerations—such as explainability, auditability, accountability, and risk mitigation—are frequently addressed only after deployment, if at all.

AI Risk, Safety, and Data Governance

AI systems introduce unique risks beyond those associated with traditional software. These include data leakage, model bias, unsafe optimization behavior, concept drift, and lack of accountability for automated decisions. In healthcare and caregiving contexts, such risks are amplified by the sensitivity of data and the vulnerability of affected populations.

Data-centric AI approaches emphasize minimizing data exposure and enforcing security controls throughout the AI lifecycle **[6]**. However, data protection alone does not address risks arising from model behavior, decision logic, or human-AI interaction.

NIST Artificial Intelligence Risk Management Framework

The NIST AI Risk Management Framework provides a structured approach to managing AI risks across organizational contexts **[3]**. The framework is organized around four core functions:

- **Govern:** Establishing accountability, policies, and oversight.
- **Map:** Understanding system context, stakeholders, and potential risks.
- **Measure:** Assessing risk likelihood and impact.
- **Manage:** Implementing mitigation strategies and monitoring effectiveness.

The AI RMF emphasizes continuous risk management rather than one-time compliance. However, it does not prescribe specific technical implementations, leaving practitioners uncertain about how to apply it in real-world systems.

Human-Centered and Trustworthy AI

Human-centered AI frameworks emphasize transparency, user agency, and meaningful human oversight **[7]**. In caregiving contexts, trust is shaped not only by accuracy but also by explainability, consistency, and perceived accountability. Caregiver-facing AI systems must therefore support contestability, override mechanisms, and shared responsibility **[10]**. In autism care, governance must also enable individualized model behavior and personalized decision support consistent with precision medicine principles, because population-level generalizations may increase risk and reduce clinical usefulness **[8]**.

Research Gap

Despite extensive discussion of AI governance, few studies demonstrate how governance frameworks such as the NIST AI RMF can be translated into **deployable technical architectures** for autism care systems. This research addresses that gap.

Governance-Centric System Architecture

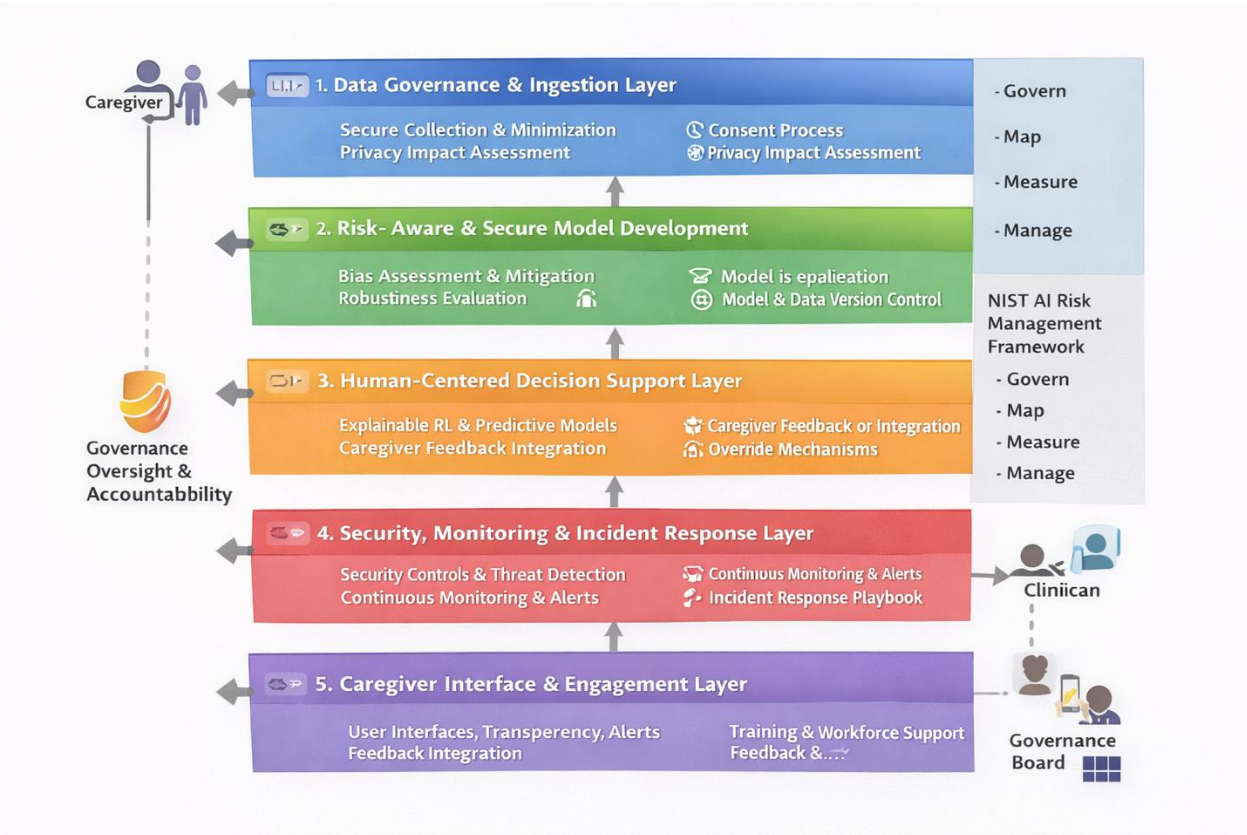
Caregiver
1. Data Governance & Ingestion Layer
Secure Collection & Minimization
Privacy Impact Assessment
Consent Process
Privacy Impact Assessment
2. Risk- Aware & Secure Model Development
Bias Assessment & Mitigation
Robustiness Evaluation
Model is epalieation
Model & Data Version Control
3. Human-Centered Decision Support Layer
Explainable RL & Predictive Models
Caregiver Feedback Integration
Caregiver Feedback or Integration
Override Mechanisms
4. Security, Monitoring & Incident Response Layer
Security Controls & Threat Detection
Continuous Monitoring & Alerts
Continiuous Monitoring & Alerts
Incident Response Playbook
5. Caregiver Interface & Engagement Layer
User Interfaces, Transperency, Alerts
Feedback Integration
Training & Workforce Support
Feedback &...
- Govern
- Map
- Measure
- Manage
NIST AI Risk Management Framework
- Govern
- Map
- Measure
- Manage
Governance Oversight & Accountabbility
Cliniican
Governance Board

Figure 1: Governance-Centric AI Architecture for Autism Care Systems

The proposed architecture consists of five integrated layers:

1. **Data Governance and Ingestion Layer**
2. **Risk-Aware Model Development Layer**
3. **Decision Governance and Explainability Layer**
4. **Monitoring, Audit, and Incident Response Layer**
5. **Caregiver and Stakeholder Interface Layer**

Each layer operationalizes one or more NIST AI RMF functions.9

Risk Taxonomy for AI-Driven Autism Care

A domain-specific risk taxonomy is essential for meaningful governance.

Data Risks

- Unauthorized data access
- Inadequate consent management
- Excessive data retention

Model Risks

- Bias across behavioral or demographic subgroups
- Concept drift due to developmental changes
- Over-optimization leading to unsafe recommendations
- Personalization risk is also relevant: failure to adapt models to individual baselines can create systematic errors, which is counter to precision AI approaches in healthcare **[8]**.

Decision Risks

- False escalation alerts
- Missed escalation events
- Inappropriate intervention recommendations

Human Interaction Risks

- Automation bias
- Caregiver overreliance or disengagement
- Misinterpretation of explanations

Operationalizing NIST AI RMF Functions

Govern: Embedding Accountability

Governance is operationalized through:

- Defined ownership of models and data pipelines
- Version-controlled documentation
- Ethical use policies embedded into system configuration

Governance artifacts are treated as system components rather sthan external documents **[3,7]**.

Map: Contextual Risk Mapping

System context is continuously mapped by:

- Identifying stakeholders (children, caregivers, clinicians)
- Mapping data flows and decision points
- Enumerating harm scenarios linked to system actions **[1,2]**
- Mapping should explicitly capture individual baseline variability and personalization requirements as first-class constraints for model development and evaluation **[8]**.

Measure: Risk Metrics and Indicators

Risk measurement extends beyond accuracy to include:

- False alert rates
- Caregiver override frequency
- Explainability satisfaction scores
- Bias indicators across profiles **[7,10]**

Manage: Risk Mitigation Controls

Risk management includes:

- Threshold-based decision gating
- Human-in-the-loop overrides
- Automated rollback of unsafe model updates
- Cybersecurity controls aligned with connected medical device risks **[6]**

Control Catalog and Audit Artifacts

A practical control catalog translates governance principles into actionable controls.

Control ID	Control Description	RMF Function
GC-01	Data minimization enforcement	Govern
GC-02	Bias monitoring across profiles	Measure
GC-03	Caregiver override logging	Manage
GC-04	Model version traceability	Govern
GC-05	Incident response workflow	Manage

Table 1: Example Governance Control Catalog

Audit artifacts are generated automatically to support internal and external review.

Implementation Blueprint

Implementation proceeds in phases:

1. Governance role definition and documentation pipelines
2. Risk-aware metric integration
3. Monitoring and audit dashboard deployment
4. Caregiver training and feedback integration

This phased approach minimizes disruption to existing care workflows.

Discussion

Embedding governance directly into system architecture shifts AI risk management from a reactive compliance exercise to a proactive design principle. For autism care systems, this approach enhances trust, accountability, and long-term sustainability **[7,10]** while mitigating cybersecurity and privacy risks **[6]**.

Limitations and Future Work

This study focuses on architectural design rather than clinical validation. Future work will evaluate governance-centric systems in real-world deployments, automate governance metrics, and align with international AI standards.

Conclusion

This research demonstrates how the NIST AI Risk Management Framework can be operationalized within AI-driven autism care systems through a governance-centric technical architecture. By translating abstract principles into concrete controls and workflows, the proposed approach advances trustworthy, auditable, and scalable AI deployment in sensitive caregiving environments **[3,7,10]**.

References

1. Islam, M. M., Hassan, M. M., Hasan, M. N., Islam, S., & Hussain, A. H. (2024). Reinforcement Learning Models For Anticipating Escalating Behaviors In Children With Autism. Journal of International Crisis and Risk Communication Research , 3225–3236. https://doi.org/10.63278/jicrcr.vi.3221

2. Islam, S., Hussain, A. H., Islam, M. M., Hassan, M. M., & Hasan, M. N. (2024). Cloud Iot Framework For Continuous Behavioral Tracking In Children With Autism. Journal of International Crisis and Risk Communication Research , 3517–3523. https://doi.org/10.63278/jicrcr.vi.3313

3. Hussain, A. H., Islam, M. M., Hassan, M. M., Hasan, M. N., & Islam, S. (2024). Operationalizing The NIST AI RMF For Smes — Top National Priority (AI Safety) And Perfect For Your Data/IT Toolkit; Produce A Lean Control Catalog, Audit Checklist, And Incident Drill For Real LLM Workflows. Journal of International Crisis and Risk Communication Research , 2555–2564. https://doi.org/10.63278/jicrcr.vi.3314

4. Hasan, M. N., Islam, S., Hussain, A. H., Islam, M. M., & Hassan, M. M. (2024). Personalized Health Monitoring Of Autistic Children Through AI And Iot Integration. Journal of International Crisis and Risk Communication Research , 358–365. https://doi.org/10.63278/jicrcr.vi.3315

5. Hassan, M. M., Hasan, M. N., Islam, S., Hussain, A. H., & Islam, M. M. (2023). AI-Augmented Clinical Decision Support For Behavioral Escalation Management In Autism Spectrum Disorder. Journal of International Crisis and Risk Communication Research, 201–208. https://doi.org/10.63278/jicrcr.vi.3312

6. 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

7. Islam, M. M., Arif, M. A. H., Hussain, A. H., Raihena, S. S., Rashaq, M., & Mariam, Q. R. (2023). Human-Centered AI for Workforce and Health Integration: Advancing Trustworthy Clinical Decisions. J Neonatal Surg, 12(1), 89-95. https://jneonatalsurg.com/index.php/jns/article/view/9123

8. Islam, M. M., & 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

9. Islam, M. M., Hussain, A. H., Mariam, Q. R., Islam, S., & 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. https://doi.org/10.57239/prn.25.03310048

10. Raihena, S. S., Arif, M. A. H., Mariam, Q. R., Hussain, A. H., & Rashaq, M. AI-Enhanced Decision Support Systems for Autism Caregivers: Redefining HR's Role in Workforce Planning and Patient-Centered Care. https://doi.org/10.63682/fhi2698