IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13202

Bridging the Intelligence Gap: A Conceptual Framework for Scaling Edge AI Across Heterogeneous Hardware

Vivek Gujar[1], Ashwani Kumar Rathore[2]

Chief Strategy Officer, Indoai Technologies P Ltd[1]

Chief Executive Officer, Indoai Technologies P Ltd[2]

Abstract: The scalability of Edge Artificial Intelligence (Edge AI) is fundamentally constrained by hardware heterogeneity and uneven compute capabilities across deployment environments. This "Intelligence Gap" represents a critical architectural barrier that prevents the democratized adoption of vision intelligence. This paper proposes a theoretical framework to address the Edge AI scalability problem through three interconnected innovations: (1) the AI Readiness Index (AIRI) for standardizing intelligence measurement; (2) a modular Appization architecture for decoupling intelligence from hardware; and (3) a Vertical Solution Stack implementing the 3Ps framework (Personalization, Platforms, Performance Analytics). Central to this framework, we introduce the EdgeBox as a "Legacy Redemption" artifact—a neural interface designed to retroactively apply the Solution Stack to non-intelligent infrastructure. This allows for the "Neural Scrubbing" of legacy systems, transforming "dumb" sensors into AIRI-certified intelligence nodes. This paper employs a design science methodology, drawing on established theories from platform economics and service-dominant logic to develop a structured approach to artifact creation. We present a comprehensive model that transforms hardware heterogeneity from a scaling constraint into a managed resource. We demonstrate how the EdgeBox mediates the transition from hardware-centric to intelligence-centric edge computing. The proposed framework provides a novel theoretical foundation for Edge AI scalability. The introduction of the EdgeBox and the 3Ps framework offers a pathway to bridge the intelligence gap, enabling modular, measurable, and trustworthy AI across heterogeneous environments and present a case study.

Keywords: Edge AI, Hardware Heterogeneity, Conceptual Framework, AI Standardization, Solution Stack, Platform Economics, Theoretical Model, Indoai, AI Cameras

I. INTRODUCTION

The Theoretical Challenge of Edge AI Scalability

Edge Artificial Intelligence represents a paradigm shift in computing architecture, promising to bring computational intelligence closer to data sources while reducing latency, bandwidth requirements, and privacy concerns[1-4]. However, the theoretical promise of Edge AI confronts a fundamental architectural constraint intrinsic to device[5]: extreme hardware heterogeneity and uneven compute capabilities across deployment environments[6]. This "Intelligence Gap" the disparity between theoretical AI capabilities and practical deployment realities across diverse hardware platforms represents a significant theoretical challenge that has not been adequately addressed in the literature [2]. Architecturally, edge devices show tremendous computational capability, from microcontrollers with limited memory and processing power to sophisticated AI accelerators with dedicated neural processing units[31]. Economically, this heterogeneity creates market failures of Akerlof's "market for lemons" [7], where information asymmetry prevents efficient procurement and deployment of AI capabilities. In the organization, the integration complexity arising from heterogeneous systems creates barriers to adoption that existing theoretical models do not adequately address.

Scalability Crisis

The premise that "Edge AI scalability is constrained by hardware heterogeneity and uneven compute capabilities" identifies the single greatest barrier to the mass adoption of intelligent vision systems. In simple terms, while cloud computing offers a uniform environment, the "Edge" is a chaotic fragmented landscape of diverse devices with vastly different "brains" (processing power).

The Problem: Hardware Heterogeneity & Uneven Compute

Scaling AI across the edge is not a simple software update because the underlying hardware varies wildly. This creates several critical bottlenecks:

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁔ Peer-reviewed & Refereed journal ⁔ Vol. 13, Issue 2, February 2026
DOI: 10.17148/IARJSET.2026.13202

- **The "Intelligence Gap":** Traditional hardware metrics (like megapixels or sensor size) do not reflect a device's actual ability[24] for example like to run complex AI models. A camera might have high resolution but lack the specialized AI chips needed for real-time inference, leading to project failure when the software outpaces the hardware.
- **Non-Uniform Architectures:** Edge devices range from low-power IoT sensors to high-performance AI cameras. A model optimized for one architecture may be completely incompatible with another, forcing developers to rewrite code for every specific device type, which is impossible to scale[26].
- **Computational Constraints:** Many edge devices have strictly limited memory and power. While a "Global Model" might work in a lab, it often fails in the field because the device cannot handle the "compute-heavy" nature of deep learning.

II. CONCEPTUAL CONTRIBUTION

This conceptual article addresses the following: *What theoretical framework could enable scalable Edge AI deployment across environments characterized by extreme hardware heterogeneity and uneven compute capabilities?*

Our primary contribution is theoretical rather than empirical:

1. **The AIRI Conceptual Framework**[8]**:** A already proposed standardization metric system for quantifying and comparing AI readiness across heterogeneous hardware.
2. **The Modular Appization Concept**[9]**:** A theoretical architecture for decoupling intelligence from hardware constraints through containerized AI applications.
3. **The Vertical Solution Stack Model:** A 'solution' stack combines technological (hardware, software, connectivity) and data nodes, with a political and economic layer shaping and simultaneously influenced by consultancy and media (Velislava Hillman)[29]. A conceptual end-to-end architecture implementing the 3Ps framework (Personalization, Platforms, Performance Analytics)[10] for orchestrating intelligence across capability gradients.
4. **Theoretical Integration:** A synthesis of concepts from platform economics, service-dominant logic and standardization theory applied to Edge AI scalability.

These conceptual contributions provide a theoretical foundation for transforming hardware heterogeneity from a scaling barrier into a managed resource, offering pathways for democratized access to Edge AI capabilities.

III. THEORETICAL FOUNDATIONS

Platform Economics and Standardization Theory

The problem of heterogeneity in technology markets has been extensively studied in platform economics literature. Shapiro and Varian [4] established that standardization is a standardization is a critical strategy to overcome market fragmentation[11] and can overcome fragmentation and create network effects in technology markets. Rietveld and Schilling [5] further developed the distinction between horizontal and vertical platform strategies[12], noting that vertical integration often proves superior in complex, regulated domains[13][14]—precisely the environments where Edge AI faces its greatest heterogeneity challenges. The standardization literature provides valuable frameworks for addressing heterogeneity. David and Greenstein [15] identify three key functions of standards: reducing variety, enabling compatibility and conveying information[15]. The AIRI framework aims to perform all three functions for Edge AI hardware, creating what we term "intelligence liquidity"—the ability to trade AI capability as a standardized commodity.

Service-Dominant Logic and Value Co-Creation

Vargo and Lusch's[16] Service-Dominant Logic (S-D Logic)[18] provides a theoretical foundation for understanding value creation in heterogeneous ecosystems. Traditional goods-dominant logic treats hardware as a value-delivery mechanism, while S-D Logic emphasizes value co-creation through service exchange among multiple actors[17]. This theoretical perspective is particularly relevant for Edge AI ecosystems, where value emerges from interactions among hardware manufacturers, software developers, system integrators and end-users.

The 3Ps framework (Personalization, Platforms, Performance Analytics) proposed in this paper operationalizes S-D Logic principles for Edge AI environments, creating theoretical pathways for continuous value co-creation across heterogeneous systems.

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 2, February 2026
DOI: 10.17148/IARJSET.2026.13202

Modular Systems Theory

Baldwin and Clark's [8] work on modularity[19] provides theoretical grounding for the Appization[9] concept. Modular systems theory suggests that complex systems can be decomposed into independently designed components that interact through standardized interfaces. This theoretical approach offers pathways for managing heterogeneity through abstraction and interface standardization.

The Appization architecture proposed in this paper applies modular systems theory to Edge AI, theoretically enabling "write once, deploy anywhere" [21][22] capabilities through standardized interfaces and abstraction layers.

IV. THE AIRI FRAMEWORK: A THEORETICAL MODEL FOR STANDARDIZING INTELLIGENCE MEASUREMENT

Conceptual Design Principles

The AI Readiness Index (AIRI)[8] is conceived as a theoretical framework based on three design principles derived from standardization literature:

1. **Multidimensional Measurement:** Following the theoretical work of Kaplan and Norton [20] on balanced scorecards, AIRI proposes a multidimensional approach that captures technical, architectural, integration and governance dimensions.
2. **Information Asymmetry Reduction:** Drawing on Akerlof's [3] work on market failures, AIRI aims to reduce information asymmetry by providing verifiable, standardized metrics for AI capability.
3. **Evolutionary Standardization:** Recognizing that AI technology evolves rapidly, the framework incorporates theoretical mechanisms for periodic revision and adaptation, similar to the evolutionary standardization processes described by Egyedi[32]. He suggests that instead of separating technology development and standards development into two separate factors, standardization should be treated as a regular environment of technology development. One important classification related to the connection between technology and standards is the distinction between designing and selecting standardization[30].

Proposed Measurement Dimensions

Theoretically, AIRI would evaluate devices across four weighted dimensions[AIRI]:

AIRI Measurement Framework

Dimension	Theoretical Weight	Proposed Metrics	Theoretical Basis
Compute & Performance	30%	Sustained TOPS, Inference Latency, Power Efficiency	Derived from computational complexity theory and performance benchmarking literature
Architecture & Flexibility	25%	OTA Update Support, Custom ModelCompatibility, Framework Support	Based on software architecture principles and backward compatibility theory
Integration & Openness	25%	Standard Protocol Compliance, API Quality, Interoperability Score	Grounded in network effects theory and compatibility standards literature
Governance & Trust	20%	Security Certification, Bias Mitigation Reporting, Data Sovereignty Controls	Derived from information governance theory and regulatory compliance literature

Theoretical Certification Model

The framework proposes a tiered certification model based on AIRI scores:

- **Certified (60-69):** Basic AI inference capability for simple applications
- **Silver (70-79):** Reliable performance for common workloads

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁓ Peer-reviewed & Refereed journal ⁓ Vol. 13, Issue 2, February 2026
DOI: 10.17148/IARJSET.2026.13202

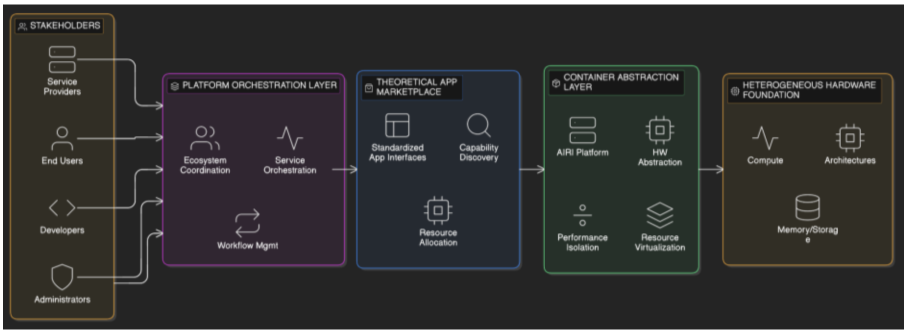
- **Gold (80-89):** High-performance with advanced features and governance
- **Platinum (90-100):** Enterprise-grade with comprehensive capabilities

The certification process would theoretically involve independent testing following established standardization processes, though the specific implementation mechanisms require further development.

V. MODULAR APPIZATION: A THEORETICAL ARCHITECTURE

Conceptual Architecture

The authors[27] explore three key design choices in modularisation: level of granularity, standardisation of interfaces, and unification of data, and their impacts on structural, dynamic, and subjective complexity The Appization architecture represents a theoretical approach to managing heterogeneity through modular design. Conceptually, it would decompose monolithic AI systems into containerized, task-specific applications with standardized interfaces[28]:



Theoretical Matching Algorithm: Conceptual App-Device Matching Algorithm

The proposed "Theoretical App-Device Matching" algorithm addresses a critical challenge in Edge AI: efficiently deploying diverse AI applications across a heterogeneous landscape of edge devices. This conceptual framework posits an intelligent orchestration layer that moves beyond simple compatibility checks, instead focusing on an optimized deployment configuration (***C***).

At its core, the algorithm evaluates each available AI application (A) against the specific **Device capability profile (*D*)** of a target edge device. This evaluation is multifaceted, assessing three key dimensions:

1. **Compatibility (*f*):** This score gauges the architectural alignment between the app's requirements (e.g., supported instruction sets, libraries) and the device's architecture (e.g., ARM, x86, NPU type).
2. **Performance (*g*):** This metric assesses how well the device's raw compute capability (e.g., TOPS, CPU cores) can meet the app's computational demands (e.g., inference speed, model size).
3. **Efficiency (*h*):** This score considers the device's power profile and the app's power requirements, crucial for battery-operated or resource-constrained edge environments.

These individual scores are then combined into a weighted total_score, using coefficients (alpha, beta, gamma) to reflect the relative importance of each dimension based on the deployment's priorities. Only applications exceeding a predefined threshold are added to a candidate_set, from which the optimal deployment configuration is ultimately selected. This algorithmic approach is fundamental to managing the "Intelligence Gap" by ensuring that applications are deployed only where they can truly deliver optimal value.

Devising a Maximization Formula

To fully formalize this conceptual algorithm, a robust mathematical formula can be devised to precisely calculate the total_score. This would involve defining the functions ***f, g***, and ***h*** with greater precision, potentially using normalized values from the AIRI framework. The ultimate goal would be to articulate an optimization problem, aiming to maximize the total score for app deployment across a fleet of heterogeneous devices, subject to various operational constraints. This ensures not just deployment, but *optimal* deployment, where resources are utilized most effectively.

VI. THE VERTICAL SOLUTION STACK

Technological Leap: The “EdgeBox” and Legacy Redemption

A critical finding in our 2026 deployments is the role of the IndoAI EdgeBox—a compact device that represents the “Great Redemption” for legacy CCTV systems.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13202

Millions of existing cameras are “dumb assets”: they record terabytes of video that is never reviewed. The EdgeBox acts as a neural interface, plugging into NVR and applying the full Solution Stack post-facto. By processing video streams locally, it enables:

- Real-time AI inference (fire detection, intrusion alerts)
- AIRI certification (achieving scores of 80+ without replacing hardware)
- OTA model updates via Neurhub™

This “Neural Scrubbing” of legacy infrastructure is vital for fragmented industries with limited CapEx budgets—schools, MSMEs, etc. It democratizes AI by making intelligence retrofit-ready, not just new-build.

Proposed Layered Architecture

The conceptual framework organizes the heterogeneity management challenge into a five-layer model:

Layer 1: Device Heterogeneity

Theoretical foundation representing diverse hardware platforms

Layer 2: AIRI Standardization

Proposed intelligence quantification and classification layer

Layer 3: Appization Modularization

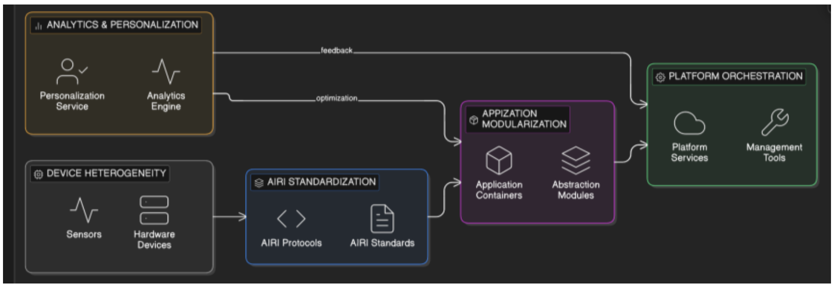
Theoretical containerization and abstraction layer

Layer 4: Platform Orchestration

Conceptual ecosystem management layer

Layer 5: Analytics & Personalization

Theoretical feedback and optimization layer



Integration with the 3Ps Framework

The conceptual Solution Stack integrates with the 3Ps framework as follows:

1. **Personalization:** Theoretical mechanisms for context-aware adaptation based on deployment environment characteristics.
2. **Platforms:** Conceptual ecosystem design connecting multiple stakeholders through standardized interfaces and governance mechanisms.
3. **Performance Analytics:** Proposed feedback loops for continuous improvement and value validation.

The 3Ps[10] are not marketing buzzwords—they are technical and strategic pillars that sustain the Solution Stack’s lifecycle:

3P Pillar	Role in Solution Stack	Technical Implementation
Personalization	Context-aware intelligence delivery	Vision AI models trained on domain data (e.g., school uniforms, warehouse pallets); user behavior triggers model recommendations[33]
Platforms	Ecosystem orchestration	Neurhub™ runtime + Appization™ marketplace; revenue sharing with devs; ONVIF/VMS compatibility[9]
Performance Analytics	Outcome validation & iteration	Real-time dashboards for accuracy, latency, TCO; federated learning for on-device refinement[34]

This triad forms a self-reinforcing flywheel: more deployments → better AIRI scores → more developer interest → richer solution packs → lower total cost of ownership (TCO).

Case Studies: The 3Ps in Practice

Education: The “Smart School Pack”

In Indian schools, the primary barrier to AI adoption was integration chaos: separate vendors for cameras, biometric logs, mobile apps, and ERP systems. IndoAI’s Smart School Pack unified these into a single stack:

DOI: 10.17148/IARJSET.2026.13202

- Personalization: Facial recognition tuned for group-selfie attendance (Indoai's Dutypar app)[40] and distress detection (crying/shouting in corridors). Models are trained on Indian student demographics to reduce bias.
- Platforms: Schools manage the entire stack via the Dutypar interface. New models—like "Exam Cheating Detection"—can be downloaded without hardware changes, thanks to Appization™[9].
- Performance Analytics: Absenteeism trends, visitor logs, and security alerts are processed on the edge, reducing bandwidth costs by up to 90% and ensuring offline operation during internet outages[26].

Result: Deployment time could be significantly reduced with teacher saving time saved on roll calls.

Retail: Hyper-Personalization in Malls[41]

Traditional digital signage is static and ignored. IndoAI's Retail Intelligence Stack turns cameras into "Engagement Engines":

- Personalization: Opted-in shoppers are recognized via facial features, and tailored ads appear on nearby displays (e.g., "You bought this shirt—see it styled!").
- Platforms: Mall owners can swap "Theft Detection v2" for "Footfall Heatmap" seasonally via the marketplace—no firmware updates needed.
- Performance Analytics: Stores track Conversion Lift—measuring exactly how many people entered after seeing a personalized ad.

Critically, all personalization includes explainability prompts ("Why am I seeing this?") and diversity controls to avoid filter bubbles—anticipating EU AI Act and India's DPDP requirements.

Conceptual Architectural Patterns[34]

The framework proposes four theoretical patterns for heterogeneous environments:

Pattern	**Architecture**	**Description**
1	Federated Learning	*Theoretical approach enabling collaborative learning across capability gradients*
2	Hierarchical Processing Model	*Conceptual distribution of workloads across compute tiers*
3	Microservices-Based Edge AI	*Theoretical decomposition of AI functions into independent services*
4	Serverless Edge Computing	*Conceptual event-driven execution model for intermittent workloads*

VII. REAL-WORLD EXAMPLES: COMPANIES FACING THESE CHALLENGES

Several global leaders and specific sectors have encountered scalability issues:

This table serves as empirical grounding for theoretical framework, demonstrating that the "Solution Stack" is a universal requirement for scaling Edge AI.

Cross-Industry Heterogeneity Analysis(sample)

Sector	**Leading Entity(for example)**	**Primary Heterogeneity Challenge**	**"Intelligence Gap" Impact**	**Current Mitigation Strategy**	**Proposed "Solution Stack" Advantage**
Automotive	**Tesla**	Disparate fleet ages; varied FSD chip versions	Older hardware cannot run current "Full Self-Driving" neural nets.[35]	**Hardware Retrofits:** Physical chip replacement (High CAPEX).	**Modular Appization:** Dynamic model pruning based on AIRI scores.
Energy	**Shell**	Extreme remote environments (Offshore) vs. Mainland Refineries.	Remote rigs may face cooling/power challenge for high-end AI accelerators.	**Lightweight FPGAs:** Specialized, rigid hardware optimization.	**EdgeBox Redemption:** Standardizing legacy rig cameras into

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13202

					the 3P framework.
Manufacturing	**Siemens**	Decades-old legacy sensors alongside modern IIoT gateways.	Modern predictive models are incompatible with 8-bit microcontrollers[36].	**Microservices:** Stripping models to minimal functional blocks.	**Neural Scrubbing:** Using the EdgeBox to process legacy data at a higher tier.
Logistics	**TCS / Amazon**	Multi-vendor telematics; diverse store/warehouse layouts.	Inconsistent latency for real-time tracking across varied network nodes[37][38].	**Contextual Shifting:** Manual tuning of models for specific sites.	**Performance Analytics:** Automated ROI validation and latency tracking.
Agriculture	**John Deere**	High-speed real-time vision on varied tractor model years.	Disconnected operation in remote fields limits cloud offloading.[39]	**Edge-Cloud Hierarchy:** Local inference for spray, cloud for logs.	**AIRI Certification:** Guaranteed sub-second latency for all certified kits.

As shown in the table above, the "Intelligence Gap" manifests in two primary ways:

1. **Temporal Heterogeneity:** The coexistence of old (legacy) and new hardware (e.g., Tesla, Siemens).
2. **Environmental Heterogeneity:** Disparities in power, cooling and connectivity (e.g., Shell, John Deere).

The Vertical Solution Stack addresses these by providing a unified orchestration layer. Rather than manually rewriting code for every tractor model or refinery rig, the AIRI layer automatically determines the "intelligence liquidity" of the node, and the Appization layer deploys the most compatible version of the model.

VIII. CASE STUDY: STANDARDIZING EDGE AI ACROSS HETEROGENEOUS INFRASTRUCTURE

The Challenge: Why Edge AI Fails at Scale

Standard AI models often break when rolled out across mixed hardware environments. Organizations frequently experience "pilot-to-rollout performance drift," where a model that works in a controlled environment fails in real-world deployment due to several critical factors:

- Hardware Fragmentation: Vast differences in CPU, GPU and NPU architectures lead to unpredictable performance.
- Hidden Bottlenecks: Video decoding loads and stream counts are often overlooked during the pilot phase, creating processing logjams in the field.
- Environmental & Operational Drift: Variables such as heat (thermal throttling), memory headroom and driver versions significantly alter performance over time.

The IndoAI Solution: The 'Compute Contract'

IndoAI solves these challenges by introducing a dedicated Edge Runtime layer between raw camera feeds and final outcomes. This runtime enforces a "compute contract" that guarantees a specific Service Level Agreement (SLA) regardless of the underlying mixed hardware.

The architecture consists of four primary internal modules:

1. Capability Profiler: Benchmarks the specific decode and inference capacity of the local device.
2. Model Registry: Maintains multi-build variants optimized for specific CPU, GPU, or NPU architectures.
3. SLA Policy Engine: Automatically tunes operational parameters like FPS, resolution, and batch sizes to meet targets.
4. Observability & Safe OTA: Provides telemetry, canary deployments, and automated rollbacks to ensure stability.

DOI: 10.17148/IARJSET.2026.13202

Operational Workflow: The Control Loop
The IndoAI Edge Runtime operates as a repeatable, continuous control loop to maintain high performance under real-site conditions:

1. Profile: Continuously benchmark the device's capacity for both decoding and inference.
2. Select: Choose the optimal model build specifically designed for the local processor (CPU, GPU, or NPU).
3. Adapt: Dynamically tune the operating point (FPS, resolution, and model tier) to hit the required SLA.
4. Verify: Monitor P95 latency, dropped frames, and thermals continuously, triggering an automatic rollback if performance degrades.

Business Outcomes: Selling Certainty, Not Models
In the IndoAI ecosystem, customers do not buy individual AI models; they purchase guaranteed outcomes. These are codified in SLA Contract Cards, which define the reliability of the rollout.
Example SLA Targets for Intrusion Detection:

- P95 Alert Latency: <=2.0 seconds.
- Minimum Usable FPS: >=12 FPS.
- Stream Support: Stable performance for up to 16 streams per node.
- Priority: Critical (Safety).

By enforcing these targets through adaptive tuning and real-time monitoring, IndoAI ensures that critical safety and operational outcomes remain stable across even the most heterogeneous edge infrastructure.

IX. CONCLUSION

This conceptual paper has presented a theoretical framework for addressing the fundamental challenge of hardware heterogeneity in Edge AI scalability. By proposing the AIRI standardization framework, the Appization modular architecture, and the Vertical Solution Stack orchestration model, we offer a structured approach to transforming heterogeneity from a constraint into a managed resource.

The framework’s contributions extend beyond Edge AI to broader questions of platform design and ecosystem orchestration in fragmented technology markets. The ultimate implication is a paradigm shift from hardware-centric to intelligence-centric computing, where AI capability becomes a standardized, tradable commodity. This shift aims to democratize AI access across socioeconomic divides, accelerating innovation across multiple sectors.

Strategic Outlook and Challenges
The Solution Stack represents a transition from selling isolated AI components to delivering certified outcomes. Enabled by the 3Ps (Personalization, Platforms, and Performance Analytics), it turns fragmented industries into programmable markets where intelligence is modular, measurable, and trustworthy.

However, the author acknowledges significant hurdles that must be addressed:

- **Governance:** Mitigating "vendor capture" of AIRI via independent bodies (e.g., ISO partnerships).
- **Standardization:** Reducing model fragmentation through universal formats like ONNX/TensorRT.
- **Ecosystem Disruption:** Supporting System Integrators (SIs) through premium certification paths.

Future Research Agenda
As AI moves beyond the cloud and into the physical world, vertical platforms will define the next era of enterprise value. Future work will expand the AIRI framework into new frontiers:

1. **Generative AI:** Integrating Large Language Models (LLMs) with a focus on "Hallucination Rates," context fidelity, and privacy preservation in domain-specific vertical applications.
2. **Physical Intelligence:** Expanding into Robotics (physical reliability/coordination) and Autonomous Vehicles (edge-case robustness).
3. **Quantum Security:** By 2030, we foresee Quantum-Safe Solution Stacks- this will explore the integration of Post-Quantum Cryptography (PQC) within the Performance Analytics pillar to ensure that “machine sight” remains private and immutable in the quantum era.

REFERENCES

[1]. Pethuru Raj, J. Akilandeswari, M. Marimuthu, Chapter Five - The edge AI paradigm: Technologies, platforms and use cases, Editor(s): Pethuru Raj, Kavita Saini, Chellammal Surianarayanan, Advances in Computers, Elsevier, Volume 127, 2022, https://doi.org/10.1016/bs.adcom.2022.02.003.

[2]. Raghubir Singh , Sukhpal Singh Gill, Internet of Things and Cyber-Physical Systems, Edge AI: A survey, Volume 3, 2023, https://doi.org/10.1016/j.iotcps.2023.02.004

[3]. Lejla Banjanovic-Mehmedovic, Anel Husaković, Edge AI: Reshaping the Future of Edge Computing with Artificial Intelligence, Conference: Basic technologies and models for implementation of Industry 4.0, October 2023, DOI: 10.5644/PI2023.209.07

[4]. Ikechukwu Innocent Umeh , Kobimdi Cordelia Umeh, Redefining the future of data processing with edge computing, World Journal of Advanced Research and Reviews, 2024, 24(02), 2865-2877, https://doi.org/10.30574/wjarr.2024.24.2.3463

[5]. Mohamad Abou Ali, Fadi Dornaika, Edge Artificial Intelligence: A Systematic Review of Evolution, Taxonomic Frameworks, and Future Horizons, arXiv:2510.01439v1 [cs.LG] 01 Oct 2025

[6]. Yue Zheng, Yuhao Chen, Bin Qian, Xiufang Shi, Yuanchao Shu, Jiming Chen, A Review on Edge Large Language Models: Design, Execution, and Applications, arXiv:2410.11845v1 [cs.DC] 29 Sep 2024

[7]. G. A. Akerlof, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, vol. 84, no. 3, pp. 488-500, Aug. 1970.

[8]. [8] Vivek Gujar, Ashwani Kr Rathore (2026). The Artificial Intelligence Readiness Index (AIRI) and the IndoAI Platform: A Novel Framework for Standardizing and Architecting Complete Vision AI Solutions., International Journal of Innovative Research In Technology (IJIRT), 12(7), 5374-5386.

[9]. Vivek Gujar, Ashwani Rathore, Appization™ @ Neurahub™: Leveraging The App Store Model For AI Functions On Edge Devices_ Case Study Of Indoai AI Camera , IOSR Journal of Computer Engineering (IOSR-JCE), Volume 26, Issue 6, Ser. 2 (Nov. – Dec. 2024)

[10]. Vivek Gujar, Marketing Mix 3.0: Personalization, Platforms and Performance Analytics in the Algorithmic Era of Growth Hacking- A Review," *International Advanced Research Journal in Science, Engineering and Technology (IARJSET)*, DOI: 10.17148/IARJSET.2025.121203

[11]. Carl Shapiro and Hal Varian, 1999 book, Information Rules: A Strategic Guide to the Network Economy,

[12]. Joost Rietveld, Melissa A. Schilling, Platform Competition: A Systematic and Interdisciplinary Review of the Literature, Journal of Management, SMA, Volume 47, Issue 6, https://doi.org/10.1177/0149206320969791

[13]. Jiatao Li, Liang Chen, Jingtao Yi, Jiye Mao, Jianwen Liao4, Ecosystem-specific advantages in international digital commerce, Journal of International Business Studies (2019) 50 (9), 1448-1463., DOI: 10.1057/s41267-019-00263-3,

[14]. Melissa A. Schilling , Technological Lockout: An Integrative Model of the Economic and Strategic Factors Driving Technology Success and Failure, Academy of Management ReviewVol. 23, Apr 1998, https://doi.org/10.5465/amr.1998.533226

[15]. Paul A. David, Shane M. Greenstein, The Economics Of Compatibility Standards: An Introduction To Recent Research, January 1990, Economics of Innovation and New Technology 1(1-2):3-41, DOI: 10.1080/10438599000000002

[16]. Vargo, S. L., & Lusch, R. F. (2016). Institutions and Axioms: An Extension and Update of Service-Dominant Logic. JAMS, 44, 5–23.

[17]. Peter Ekman, Randle D. Raggio, Steven M. Thompson , Service network value co-creation: Defining the roles of the generic actor, Industrial Marketing Management Volume 56, July 2016, Pages 51-62, https://doi.org/10.1016/j.indmarman.2016.03.002

[18]. https://www.ebsco.com/research-starters/social-sciences-and-humanities/service-dominant-s-d-logic

[19]. Stefano Brusoni, Joachim Henkel, Michael G. Jacobides , The power of modularity today: 20 years of "Design Rules", January 2023, Industrial and Corporate Change 32(1), DOI: 10.1093/icc/dtac054

[20]. Kaplan, Robert S., David Norton. "The Balanced Scorecard: Measures that Drive Performance." Harvard Business Review 70, no. 1 (January–February 1992): 71–79. (Reprint #92105.)

[21]. Sagar Choudhary, Vijitha Sriramulu, Edge AI Deploying Artificial Intelligence Models on Edge Devices for Real-Time Analytics, ITM Web of Conferences 76:01009, March 2025, DOI: 10.1051/itmconf/20257601009

[22]. https://www.edgeaifoundation.org/livestreams/edge-ai-talks-write-once-deploy-everywhere-the-model-agnostic-ai-deployment

[23]. Henk J. de Vries, Hugo Verheul, Harmen Willemse, Stakeholder identification in IT standardization processes, December 2003, In book: Proceedings of the Workshop on Standard Making: A Critical Research Frontier for Information Systems

[24]. Haoyang Wang, Ruishan Guo, Pengtao Ma, Ciyu Ruan, Xinyu Luo, Wenhua Ding, Tianyang Zhong, Jingao Xu, Yunhao Liu, Xinlei Chen , Towards Mobile Sensing with Event Cameras on High-mobility Resource-constrained Devices: A Survey, arXiv:2503.22943v1 [cs.RO] 29 Mar 2025

[25]. M.Z. Nase , When machine learning models retire, decay, or become obsolete: A review on algorithms, software, and hardware, Renewable and Sustainable Energy Reviews, Volume 226, Part A, January 2026, 116231, https://doi.org/10.1016/j.rser.2025.116231

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ꕥ Peer-reviewed & Refereed journal ꕥ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13202

[26]. Massimo Merenda, Carlo Porcaro, Demetrio Iero , Edge Machine Learning for AI-Enabled IoT Devices: A Review, Sensors 2020, 20(9), 2533; https://doi.org/10.3390/s20092533

[27]. van den Hooff, B., Jochemsen, E., Rezazade Mehrizi, M., Plomp, M., & Mol, K. (2026). Modularisation and the management of IT architecture complexity. European Journal of Information Systems, 35(1), 90–108. https://doi.org/10.1080/0960085X.2025.2546388

[28]. Yalemisew Abgaz, Andrew McCarren, Peter Elger, David Solan, Neil Lapuz, Marin Bivol, Glenn Jackson, Murat Yilmaz, Jim Buckley, Paul Clarke; Decomposition of Monolith Applications Into Microservices Architectures: A Systematic Review, IEEE Transactions on Software Engineering, Volume 49, Issue 8, https://doi.org/10.1109/TSE.2023.3287297

[29]. Velislava Hillman· Molly Esquivel , The 'Solution Stack' of a Neoliberal Inferno Apparatus: A Call for Teacher Conscience, https://doi.org/10.1007/s42438-023-00442-4 , Springer 2023

[30]. Mikael Laakso , Arto Kiviniemi , THE IFC STANDARD - A REVIEW OF HISTORY, DEVELOPMENT AND STANDARDIZATION Journal of Information Technology in Construction, http://www.itcon.org/2012/9

[31]. Vivek Gujar (2025). The Integration of Artificial Intelligence into Edge Camera Systems: A Perspective through Disruptive Innovation Theory_Case Study of Indoai., International Journal of Innovative Research In Technology (IJIRT), 12(4), 613-619.

[32]. T. M. Egyedi, Shaping Standardization: A Study of Standards Processes and Standards Policies in the Field of Telematic Services," PhD Thesis, Delft University of Technology, 1996.

[33]. Vivek Gujar, Vision LLMs: Bridging Language and Visual Understanding - Case study of IndoAI, April 2025, International Journal of Science and Research (IJSR) 14(4):158-164, DOI: 10.21275/SR25331222304

[34]. Vivek Gujar, Smarter at the Edge: Evaluating Decentralized AI Deployment Models in Federated, Hierarchical, Microservices and Serverless edge AI Architectures., June 2025, Scientific Computing and Instrumentation 14(6):1-16

[35]. https://www.autonews.com/article/20180827/OEM01/180829843/tesla-prefers-its-self-driving-hardware-over-that-of-other-makers/

[36]. https://medium.com/data-science/ultra-tinyml-machine-learning-for-8-bit-microcontroller-9ec8f7c8dd12

[37]. Diana Andreea Popescu, Noa Zilberman, Andrew W. Moore , Characterizing the impact of network latency on cloud-based applications' performance Diana Andreea Popescu, Technical reports published by the University of Cambridge Computer Laboratory, Nov 2017, http://www.cl.cam.ac.uk/techreports

[38]. Amanda Carter et al., Impact of Network Latency on Cross-Platform Database Performance, May 2024

[39]. Kiryong Ha, Grace Lewis† , Soumya Simanta† , Mahadev Satyanarayanan , Cloud Offload in Hostile Environments , School of Computer Science Carnegie, Mellon University, December 2011 CMU-CS-11-146

[40]. Vivek Gujar, Attendance Tracking And Analytics: Embracing Facial Recognition Technology In Corporate And Educational Institutions – Case Study_Dutypar, November 2023, International Research Journal of Modernization in Engineering Technology and Science 5(11):108-116, DOI: 10.56726/IRJMETS45892

[41]. Vivek Gujar, AI Camera: Unique Technique to Hyper AI Personalization for Malls_IndoAI Case Study, April 2024, International Journal of Science and Research (IJSR) 13(4):455-460, DOI: 10.21275/SR24406112210