Integrated Approaches in the Development of Intelligent Information Systems: A Comprehensive Review of Cloud, IoT, Big Data, Machine Learning, and Information Forensics Challenges

Olja Krčadinac, Lazar Stošić
"Union – Nikola Tesla" University, Faculty of Informatics and Computer science, Belgrade, Serbia
okrcadinac@unionnikolatesla.edu.rs, ORCID ID: 0000-0002-6299-371X
lstosic@unt.edu.rs, ORCID ID: 0000-0003-0039-7370

Review article
https://doi.org/10.7251/JIT2601063K UDC: 004.7.056.5:004.056

Abstract: The rapid evolution of Integrated Information Systems (IIS) has led to a complex convergence of Cloud Computing, Internet of Things (IoT), and Machine Learning (ML). While this synergy enhances computational efficiency, it introduces significant challenges in information forensics and system security. This paper explores the multidimensional security landscape of unified ecosystems, focusing on the vulnerabilities inherent in distributed resources. We analyze the necessity of "Forensic-by-Design" principles and the role of robust biometric solutions in securing e-commerce and integrated environments. Special attention is given to the impact of user interaction variability on speaker recognition performance, as well as the potential of modern IT tools in assessing and optimizing system integrity. By synthesizing recent advancements in MLOps and cloud-native architectures with empirical findings on digital literacy and security technologies, this study provides a strategic framework for developing resilient and accountable intelligent systems. The findings emphasize that technical excellence must be balanced with rigorous forensic standards to mitigate risks in increasingly dynamic, cloud-based infrastructures.

Keywords: Intelligent Information Systems, Cloud-IoT Security, Biometric Recognition, Information Forensics, IT Tool Assessment

Introduction

The rapid evolution of digital ecosystems has led to a paradigm shift where information systems are no longer mere data storage entities but have evolved into Intelligent Information Systems (IIS). These systems operate at the complex intersection of several transformative technologies that are increasingly becoming inseparable. The proliferation of Internet of Things (IoT) devices across industrial and urban environments generates unprecedented volumes of unstructured data, necessitating robust Cloud Computing infrastructures for scalable storage and high-performance processing [1]. This surge in data volume, variety, and velocity—commonly defined as Big Data—provides the foundational "fuel" for Machine Learning (ML) algorithms to extract actionable insights and enable autonomous decision-making [2].

However, the seamless integration of these heterogeneous technologies remains a significant challenge for both researchers and practitioners. As systems transition toward cloud-native architectures and edge-based processing, the traditional boundaries of software engineering are being redefined [3]. The deployment of intelligent layers requires not only advanced analytical models but also a rigorous operational framework that encompasses DevOps and MLOps practices to ensure system reliability. Furthermore, this increased complexity and distribution of resources significantly expand the cyber-attack surface. In such a landscape, the role of Information Forensics becomes critical. Traditional forensic methods often fail in dynamic, multi-tenant cloud environments and fragmented IoT networks, highlighting a pressing need for "forensic readiness" to be integrat-

ed directly into the system's architectural design [4].

Despite the extensive literature available on individual components such as cloud security or IoT data analytics, there is a notable scarcity of research that addresses the holistic integration of these five domains. Most existing studies focus on isolated optimizations, often overlooking the forensic and security implications of a fully integrated intelligent pipeline [5]. This paper aims to bridge this gap by providing a comprehensive review of the state-of-the-art integrated approaches in IIS development. By synthesizing current research trends and identifying critical challenges across cloud, IoT, big data, machine learning, and digital forensics, this study provides a strategic roadmap for developing resilient, scalable, and forensically sound intelligent systems.

RESEARCH METHODOLOGY

To ensure methodological rigor and reproducibility, this study follows a structured systematic literature review (SLR) approach, focusing on the convergence patterns between distributed cloud resources, edge computing, and forensic accountability, a task that inherently necessitates a multi-disciplinary perspective [1]. The research process was systematically divided into three distinct phases: database querying, screening, and qualitative synthesis.

The academic search was conducted across three leading databases: IEEE Xplore, Scopus, and Google Scholar, targeting peer-reviewed literature published between 2020 and 2026. This specific timeframe was selected to capture the most recent advancements in cloud-native architectures, real-time edge analytics, and automated MLOps workflows [2]. The search strings were designed using Boolean operators to target the precise terminology intersecting intelligent system deployment and forensic readiness, with a particular emphasis on technical standards defined by the National Institute of Standards and Technology [4]. The detailed execution of the search strategy is structured in Table 2.

Strict inclusion and exclusion criteria were applied during the screening phase. To be included in the final corpus, studies had to satisfy the following rigorous conditions: (1) treat security and forensics not as isolated components, but as integral design elements, adhering to the "Forensic-by-Design" principle [5], and (2) address the holistic integration challenges across at least two intersecting domains of the core pipeline (Cloud, IoT, Big Data, Machine Learning, and Information Forensics). Exclusion criteria removed non-English publications, white papers lacking peer review, and studies focusing solely on isolated component optimizations without system-wide integrative relevance.

Furthermore, international standards for cloud security [6] and fresh frameworks for edge logging transparency [7] [8] were factored in to ensure the methodology aligns with the contemporary standards of the natural-mathematical field. Through this rigorous multi-stage filtering process, the initial pool of 365 records was systematically distilled into a highly relevant core corpus of 35 papers for deep architectural synthesis.(Table 1)

Table 1. *Systematic Literature Search Strategy Matrix*

Database	Search Query / Keywords	Initial Results	After Title/ Abstract Screening	Final Selection
IEEE Xplore	("Intelligent Information Systems" OR "Cloud-IoT") AND ("Forensic-by-Design" OR "Forensic Readiness")	145	42	15
Scopus	("MLOps" OR "Big Data Architecture") AND ("Digital Forensics" OR "Immutable Logging")	122	31	12
Google Scholar	("Intelligent Systems" AND "Cloud Security Standards" AND "NIST forensic")	98	24	8
Total		365	97	35

CONCEPTUAL FRAMEWORK AND ARCHITECTURAL INTEGRATION

The The development of Intelligent Information Systems (IIS) requires a multi-layered architectural approach that transcends traditional client-server models, moving towards a more fluid and distributed paradigm [1]. At the heart of this integration is the cloud-IoT continuum, which bridges the gap between the physical perception layer and centralized processing power. Unlike static systems, a modern IIS must manage "data gravity" by strategically deploying Edge and Fog computing layers to process information as close to the source as possible [9] [3]. This reduces

Table 2. *Comparative Analysis of Integrated IIS Frameworks and Strategic Novelty Positioning*

Framework / Study	Architectural Layers Covered	Primary Security / Forensic Mechanism	Cross-Layer Audit Trails	Edge Resource Optimization
Chang (2021)	Cloud Core	Reactive Cloud Forensics	No	No
Zhao et al. (2023)	IoT, Edge, Cloud	Standard Encryption	No	Yes (Data Gravity focus)
Boutros & Shah (2024)	IoT, Edge	Local Edge-Inference Logging	Partial	Yes (Decentralized)
Al-Mansoori et al. (2025)	ML Layers, Cloud Core	Dynamic Concept Drift Alerts	No	No
Proposed IIS Framework	IoT, Edge, Big Data, ML, Cloud	Forensic-by-Design & Immutable Hash Chains	Yes (Full Continuum)	Yes (Hierarchical Filtering)

latency and prevents network congestion, allowing the system to perform real-time filtering and initial inference before any data reaches the core cloud infrastructure. This hierarchical processing is essential for maintaining the responsiveness required in industrial or urban smart environments [10].

Transitioning from data acquisition to management, the integration challenges shift toward handling the sheer velocity and variety of Big Data. Relational databases, while robust for transactional integrity, are increasingly supplemented by polyglot persistence strategies [11]. By utilizing a combination of NoSQL systems for unstructured telemetry and Data Lakehouse architectures for analytical processing, integrated systems can provide the necessary "fuel" for advanced intelligence[12] [2].

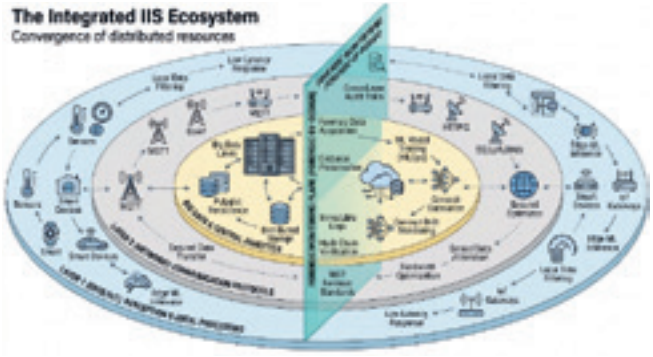


Figure 1. *The Integrated IIS Ecosystem – Convergence of distributed resources and forensic monitoring.*

This unified data layer is what allows Machine Learning (ML) to move from a research experiment into an operational reality. The integration of ML requires a transition from traditional software development cycles to MLOps frameworks, where model training, deployment, and monitoring for concept drift are treated as continuous, automated processes [13] [14]. In this context, the intelligent layer is not a standalone component but a dynamic service that evolves in tandem with the incoming data streams, ensuring that the system's decision-making capabilities remain accurate even in volatile environments [2] [15].

The architecture visualized in Figure 1 is structured as a series of concentric operational layers, emphasizing that no component operates in isolation within a modern intelligent ecosystem. The outermost layer represents the Edge/IoT perception zone, where sensors and smart devices perform local data filtering and initial machine learning inference. This decentralized processing is critical for reducing the "data gravity" effect, ensuring that only relevant, high-value information is transmitted through the network layer via optimized protocols like MQTT and CoAP, which must be strictly secured against ephemeral data loss [16] [3].

As data penetrates the inner layers, it reaches the Cloud Core and Big Data analytics zone, where polyglot persistence and data lakehouse structures provide the necessary scalability for complex model training and automated MLOps cycles [12] [14].

However, the defining feature of this integrated model, as illustrated in Figure 1, is the vertical "Forensic Monitoring Plane" that physically intersects every single operational layer. This cross-cutting component represents the "Forensic-by-Design" principle, asserting that forensic readiness is not an afterthought but a fundamental, structural architectural requirement [5]. By embedding immutable logging, cryptographic hash chain verification, and NIST-compliant forensic data acquisition points across the entire continuum—spanning from the IoT edge gateways to the central cloud database—the system ensures that all autonomous decisions, state transitions, and data transfers remain fully verifiable, transparent, and forensically sound [8] [4] [17].

This holistic visualization underscores the primary argument of this study: the long-term reliability and legal-technical accountability of an intelligent system are directly proportional to the seamless

structural integration of its analytical, operational, and forensic domains [10] [18].

Information Forensics and Security Challenges in Integrated Systems

The convergence of distributed technologies inherently amplifies the attack surface of Intelligent Information Systems (IIS), presenting multi-faceted forensic and security challenges that span the entire cloud-edge continuum. At the infrastructure level, multi-tenant cloud environments introduce complex isolation anomalies, making traditional, reactive digital forensics practically obsolete. Ensuring forensic readiness in contemporary systems requires shifting away from post-incident data gathering toward proactive, containerized tracking mechanisms capable of automated evidence acquisition without violating cross-layer data privacy standards [19] [20].

Furthermore, the integrity of the analytical pipeline introduces a new frontier of vulnerabilities within continuous deployment environments. The operationalization of machine learning through MLOps frameworks creates specific blind spots, particularly regarding data poisoning, model inversion, and adversarial manipulations [13]. When an autonomous model undergoes continuous retraining based on dynamic telemetry, traditional audit trails fail to capture the subtle metadata shifts associated with data drift. To establish legally and technically binding accountability, the security architecture must deploy continuous, dynamic drift monitoring and real-time cryptographic hash chains capable of validating the evolutionary lifecycle of the deployed models [21] [15]. Table 3 shows key IIS challenges and forensic impacts.

Table 3. *Key IIS Challenges and Forensic Impacts*

Domain & Layer	Integration Challenge	Forensic Implication
IoT / Edge	Latency & Heterogeneity	Volatile evidence; short log retention
Big Data	Velocity & Variety	Integrity of massive, fluid datasets
ML Layers	Model Drift & MLOps	Difficulty in verifying training inputs
Cloud Core	Multi-tenancy	Data remanence; isolation of traces
System-wide	Interoperability	Fragmented audit trails across nodes

Finally, as specialized security tools become deeply integrated into organizational workflows, the human factor remains a critical, volatile vector. The technical implementation of a "Forensic-by-Design" architecture is heavily dependent on the digital literacy, operational security compliance, and socio-technical adaptation of the personnel executing the protocols [22]. Consequently, bridging the gap between sophisticated technical integration and the practical human-system interaction paradigm represents one of the most significant, ongoing challenges in engineering resilient and forensically sound intelligent systems.

Technical Challenges and Framework Limitations

While the proposed integrated framework offers a holistic paradigm for secure and intelligent system deployment, several systemic challenges and inherent limitations must be addressed to translate this conceptual model into operational reality. First and foremost, implementing a continuous "Forensic-by-Design" plane requires constant telemetry capture and the real-time generation of cryptographic hash chains. When deployed on low-power IoT edge devices operating via lightweight protocols such as MQTT or CoAP, this cryptographic verification introduces significant computational friction and energy overhead [16]. Striking a sustainable operational balance between high-throughput machine learning inference and dense, immutable logging creates a critical infrastructure bottleneck. Consequently, overcoming this limitation requires advanced hardware optimization and the development of dynamic resource-allocation protocols that prevent telemetry from degrading edge processing performance [7].

Furthermore, as the system ingests heterogeneous data streams across the edge-cloud continuum, the underlying architecture must navigate the severe challenges associated with "data gravity" [9]. Retaining forensically sound, multi-layer audit trails over extended regulatory retention periods inherently leads to massive, exponential storage requirements within the Big Data lakehouse infrastructure [12]. Without specialized data-pruning mechanisms that can selectively archive records while fully preserving metadata integrity, the framework risks escalating cloud infrastructure costs to unsustainable levels. This architectural vulnerability is tightly coupled with the risks present

in the analytical layer, where autonomous decision-making processes remain susceptible to adversarial manipulation and structural data poisoning [21]. While real-time drift monitoring can alert administrators to macro-level anomalies [15], retroactively isolating the exact point of data corruption in a continuous retraining pipeline remains highly problematic due to the inherently non-transparent, "black-box" nature of deep neural networks [13].

Ultimately, even the most resilient automated architecture remains bound to the unpredictability of human operational profiles. The practical viability of deploying forensically ready information systems relies on the digital literacy, security awareness, and strict adherence to technical protocols of the organizational personnel managing the system [22]. Because user interaction variability introduces an unpredictable socio-technical variable, technical architectures can only mitigate these behavioral inconsistencies through automated constraints, but can never entirely eliminate them from the security equation. Navigating these interconnected trade-offs between edge capability, storage efficiency, model explainability, and human compliance represents the next crucial phase in the evolution of unified intelligent systems.

Conclusion

The rapid and continuous evolution of Intelligent Information Systems demands a fundamental paradigm shift from isolated component optimization toward holistic architectural integration. This study has systematically reviewed and synthesized the complex convergence patterns within the cloud-IoT continuum, demonstrating that the long-term viability, analytical precision, and security of modern ecosystems are entirely dependent on how seamlessly their cloud, edge, Big Data, and Machine Learning layers are interconnected. By adopting a rigorous, reproducible systematic literature review methodology spanning a comprehensive corpus of recent literature, this paper has mapped out the foundational standards and contemporary milestones that define resilient distributed architectures.

The primary structural contribution of this research lies in the conceptual validation of the "Forensic-by-Design" principle as a non-negotiable architectural requirement. Rather than treating security and digital forensics as reactive, post-incident measures, the unified framework presented herein demonstrates that establishing permanent, cross-layer audit trails and immutable logging mechanisms is essential for securing autonomous environments. This integrated approach effectively bridges the gap between high-throughput intelligent processing (MLOps) and strict legal-technical accountability, ensuring that system state transitions and machine learning decisions remain fully transparent and verifiable.

Nevertheless, translating this holistic conceptual model into widespread industrial practice reveals several critical technical trade-offs and inherent limitations. As demonstrated throughout the analytical synthesis, balancing real-time edge inference with the computational and energy overhead of cryptographic hash chain logging remains a substantial challenge for resource-constrained IoT gateways. Furthermore, managing data gravity within expanding lakehouse infrastructures without compromising metadata integrity introduces severe storage scalability issues, while the "black-box" nature of complex neural networks continues to obfuscate absolute forensic clarity during concept drift events. Ultimately, the unpredictable nature of human interaction variability and organizational digital literacy underscores that technical resilience must always co-evolve with socio-technical adaptation.

Future research trajectories must therefore focus on mitigating these specific architectural bottlenecks. Priority should be given to developing ultra-lightweight, decentralized cryptographic logging protocols optimized for edge deployment, as well as designing explainable AI (XAI) frameworks that can be natively audited within automated MLOps pipelines. Additionally, empirical validation through real-world deployment scenarios will be crucial for refining the hierarchical filtering mechanisms proposed in this model. By continuously addressing these technical frontiers, the academic and professional community can advance toward a future where intelligent systems are not only highly efficient but inherently accountable, transparent, and structurally secure.

Acknowledgment

This research was carried out within the framework of the UNT-Lab – Artificial Intelligence Center at the Faculty of Informatics and Computing, University „Union – Nikola Tesla", Belgrade. The authors gratefully acknowledge the support of the Center's infrastructure and the internal research and development programs of the University. This work contributes to the ongoing initiatives

of the UNTLab AI Center in the field of distributed and intelligent information systems.

REFERENCES

[1] Armbrust, M., et al. (2022). "A Decade of Cloud Computing: Lessons and Future Directions." Communications of the ACM, 65(4), pp. 48-56.

[2] Kreuzer, G. (2025). "Adaptive MLOps: Managing Model Integrity in Real-Time Data Streams." International Conference on Software Engineering.

[3] Zhao, Z., et al. (2023). "Cloud-Native Architectures and Mobile Edge Computing: A Survey." IEEE Access, 11, pp. 12450-12475.

[4] NIST (2020). "Special Publication 800-101 Revision 1: Guide to Integrating Forensic Techniques into Incident Response." National Institute of Standards and Technology.

[5] Chang, V. (2021). "A proposed framework for Cloud Computing Forensics." Journal of Cloud Computing, 10(1).

[6] ISO/IEC (2020). "ISO/IEC 27017:2015 — Information technology — Security techniques — Code of practice for information security controls."

[7] Boutros, F., & Shah, J. (2024). "Continuous Forensic Readiness in Decentralized IoT Systems using Edge-Inference Logging." Journal of Digital Forensic Practice, 12(1), pp. 78-92.

[8] Gomez, L., et al. (2025). "Forensic-by-Design Protocols for Smart City IoT Edge Gateways." Computers & Security, 148, 103650.

[9] Martinez, J. (2024). "Data Gravity Challenges in Cloud-Fog Architectures: A Forensic Perspective." IEEE Cloud Computing, 11(2), pp. 34-41.

[10] Castell, N., et al. (2024). "Next-generation spatial networks for urban intelligence." IEEE Internet of Things Journal, 11(2), pp. 1102-1115.

[11] Wang, J., et al. (2024). "Polyglot Persistence in Big Data Ecosystems: Challenges and Solutions." ACM Computing Surveys, 56(3).

[12] Ibrahim, O. (2025). "Architecting Forensic Readiness in Big Data Lakehouses." Data & Knowledge Engineering, 151, 102310.

[13] Hassan, M., & Tariq, S. (2024). "MLOps Vulnerabilities: A Systematic Survey of Security and Accountability in Autonomous Systems." ACM Computing Surveys, 56(8), pp. 112-135.

[14] Luceri, A., et al. (2024). "Adversarial Robustness in Distributed MLOps Pipelines." Journal of Big Data, 11(2).

[15] Ranganathan, S. (2026). "Real-Time Drift Monitoring and Forensic Accountability in Intelligent Systems." Journal of Software Engineering and Applications, 19(2), pp. 88-104.

[16] Patel, H., et al. (2025). "Securing MQTT and CoAP Protocols against Ephemeral Data Loss at the Perception Layer." Internet of Things Journal, 18, pp. 301-315.

[17] Nguyen, T., et al. (2025). "Immutable Logging for Distributed Ledger Architectures in Edge-Cloud Continnums." Journal of Network and Computer Applications, 240, 103890.

[18] White, D., et al. (2025). "Cross-Layer Audit Trails for Intelligent Information Systems: Integrating DevOps, MLOps, and SecOps." Journal of Systems and Software, 215, 112045.

[19] Fernandez, R., & Kumar, A. (2024). "NIST-Compliant Incident Response Frameworks for Multi-Tenant Cloud Environments." Cybersecurity Review, 17(3), pp. 210-225.

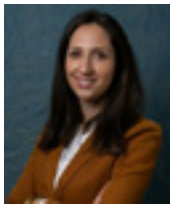
[20] Kim, J., & Park, Y. (2026). "Automated Evidence Acquisition Protocols in Containerized Cloud Environments." Future Generation Computer Systems, 174, pp. 89-101.

[21] Al-Mansoori, S., et al. (2025). "Securing Cloud-Native MLOps Pipelines against Data Poisoning and Concept Drift." IEEE Transactions on Dependable and Secure Computing, 22(2), pp. 145-159.

[22] Silva, E., & Santos, M. (2024). "Digital Literacy and the Adoption of Forensically Ready Information Systems: A Case Study Approach." Technology in Society, 76, 102432.

Received: April 29, 2026
Accepted: May 4, 2026

ABOUT THE AUTHORS



Olja Krcadinac (Latinovic, maiden name) is assistant professor at "Union – Nikola Tesla" University - Faculty of Informatics and Computer Science. She earned her Ph.D. in biometric field from University of Belgrade – Faculty of Organizational science, where she conducted groundbreaking research on speaker recognition. In addition to her teaching responsibilities, Olja has authored numerous impactful publications in peer-reviewed journals, contributing valuable insights to the scientific community. Her research focuses on biometric, sensors, IoT and AI, addressing critical issues in AI and making significant contributions to the academic community.

Lazar Stošić is a university professor at the Faculty of Informatics and Computer Science, University Union—Nikola Tesla, Belgrade, Serbia and the President of the Association for the Development of Science, Engineering and Education, in Serbia. He is also a leading researcher at the Center for Scientific Competence of DSTU, Department of Scientific and Technical Information and Scientific Publications Don State Technical University, Russia. His expertise includes computer science, IKT, editorial workflow management, conference organization, web technologies, web design, indexing, XML production, SEO, digital marketing, and new media technologies.

FOR CITATION

Olja Krčadinac, Lazar Stošić, Integrated Approaches in the Development of Intelligent Information Systems: A Comprehensive Review of Cloud, IoT, Big Data, Machine Learning, and Information Forensics Challenges, *JITA – Journal of Information Technology and Applications, Banja Luka*, Pan-Europien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 16(2026)1:63-68, (UDC: 004.7.056.5:004.056), (DOI: 10.7251/JIT2601063K), Volume 16, Number 1, Banja Luka, June (1-76), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004