Sarcouncil Journal of Engineering and Computer Sciences

ISSN(Online): 2945-3585

Volume- 04| Issue- 07| 2025

Research Article **Received:** 13-06-2025| **Accepted:** 07-07-2025 | **Published:** 26-07-2025

AI-Integrated Microservices: Advancing Decision Intelligence in Compliance-Driven Cloud Architectures

Manisha Ponugoti
Wright State University, USA

Abstract: This article explores the integration of Artificial Intelligence within microservices architectures for regulated industries, proposing a framework that balances intelligent decision-making with compliance requirements. It examines how distributed architectures create opportunities for embedding AI directly within service components while maintaining regulatory alignment. The architectural framework leverages event-driven patterns, containerization, and service mesh technologies to enable real-time inference with appropriate governance. Key compliance mechanisms, including model explainability, traceability through versioned inputs/outputs, privacy regulation alignment, and a comprehensive auditing framework, are detailed. Through case studies in clinical trials, insurance underwriting, and financial fraud detection, the article identifies common implementation patterns that support regulatory requirements across domains. Looking forward, it anticipates emerging standards for AI governance in distributed systems, evolving developer roles that span technical and ethical concerns, research opportunities in explainable AI for microservices, and technical roadmaps for compliance-aware decision intelligence. The convergence of microservices and AI represents a maturing engineering landscape where systems deliver both intelligence and accountability.

Keywords: Decision intelligence, Compliance-driven architecture, AI governance, Explainable microservices, Regulatory technology.

INTRODUCTION

The landscape of corporate computing systems has profoundly shifted during the last decade, transitioning from unified applications toward fragmented microservices that deliver remarkable scaling capabilities and deployment adaptability. This change signifies more than just technical preference—it mirrors fundamental organizational adaptations responding to market necessities for swift iteration and system durability. The microservices methodology approaches software creation as collections of compact services, each operating independently and interacting through streamlined protocols, frequently using HTTP-based resource interfaces. These individual services concentrate on particular business functions, deploy separately via automated systems, demand minimal centralized oversight, and may employ varied coding languages and information storage technologies (Lewis, J., & Fowler, M. 2014). This architectural progression has facilitated continuous deployment practices while permitting different system elements to progress at varying speeds according to business imperatives.

As these fragmented architectures evolve, they generate possibilities for incorporating intelligence precisely where information and operational logic connect. Microservices appeared specifically to address shortcomings in conventional designs, where adjustments to any component demanded complete application redeployment, and expansion occurred at the system level rather than for separate functions. By dividing applications into distinct services with dedicated information repositories and clear communication boundaries, organizations benefit from deployment versatility, technology variation, and structural alignment with business domains (Lewis, J., & Fowler, M. 2014). This framework enables groups to function simultaneously on independent business capabilities with reduced coordination requirements, achieving development momentum impossible within monolithic structures.

The expansion of microservices unveils a novel territory: embedding intelligence throughout distributed frameworks. While microservices effectively handle operational concerns, including scalability and deployment flexibility, they traditionally depend on fixed logic rather than adaptive decision processes. This constraint becomes evident as organizations attempt to automate intricate workflows requiring subtle judgments previously reserved for human specialists. Incorporating AI directly within microservices represents a logical advancement, converting these components from basic executors of established instructions into intelligent entities capable of categorization, forecasting, and irregularity identification during live operations.

This integration faces unique barriers in regulated industries as healthcare, banking, and insurance. There are existing government, corporate, academic, and public ethical frameworks for AI that share core principles of transparency, fairness, safety, accountability, and privacy, creating a

***Corresponding Author:** Manisha Ponugoti

DOI- https://doi.org/10.5281/zenodo.16452740

normative basis for ethical uses of AI (Jobin, A. *et al*., 2019). Medical research requires patient information processing that follows privacy regulations while delivering transparent reasoning for eligibility determinations. Insurance assessment utilizing AI must demonstrate fairness and consistency to satisfy industry regulations. Throughout these sectors, common requirements surface: decisions must create audit pathways, models must offer explanations, and processing must honor jurisdictional data handling requirements.

Recent industry analyses reveal compelling adoption metrics across these sectors: healthcare organizations implementing AI-integrated microservices report a 38% adoption rate with accompanying 27% reduction in process time; financial institutions demonstrate 64% adoption with 41% improvement in decision accuracy; while insurance providers show 42% adoption resulting in 31% cost reduction across underwriting processes.

This article suggests that AI-integrated microservices present a viable architectural pattern for delivering instantaneous decision intelligence while preserving regulatory alignment. Organizations can implement complex decisions automatically without losing traceability by embedding machine learning in event-driven microservices. The methods used follow ethical best practices of AI principles across the world and benefit from the deployability, diversity of technology options, and organisational alignment of forward-thinking, fit-for-purpose business microservices. The outcome is a new class of applications where intelligence is distributed throughout the system rather than centralized, enabling contextual decision-making that honors both local constraints and global governance frameworks.

ARCHITECTURAL FRAMEWORK FOR AI-INTEGRATED MICROSERVICES

Incorporating artificial intelligence within microservices frameworks requires thoughtful architectural planning that considers both distributed application characteristics and machine learning workflow demands. This segment explores foundational design patterns enabling AI-integrated microservices to provide instantaneous decision capabilities while preserving the expandability and robustness necessary in production settings.

Event-driven architecture forms the structural foundation for intelligent workflow mechanization in microservices environments. Unlike conventional request-response interactions, event-driven frameworks separate services through asynchronous messaging, permitting AI components to analyze information flows without impeding operational processes. This architectural strategy proves especially beneficial for compliance-oriented domains requiring traceable and verifiable decisions. By structuring business activities as event sequences rather than direct service interactions, organizations establish natural integration points for AI while maintaining thorough event documentation satisfying regulatory obligations. Event sourcing methodologies preserve comprehensive state change histories, creating unalterable records that support both compliance auditing and machine learning model training. Combining event-driven structures with domain-driven design principles enables teams to establish bounded contexts where AI services evolve independently while honoring domain limitations. These events travel through the messaging infrastructure, generating persistent streams that AI services process without compromising system responsiveness (Newman, S. 2021).

Containerization via technologies such as Docker delivers the separation and transferability essential for AI workloads, while Kubernetes supplies orchestration functionality to administer these containers across distributed environments. This pairing addresses several AI-specific challenges: varied resource requirements, model versioning, and reproducible environment necessities. Kubernetes capabilities extend beyond basic container management to offer sophisticated scheduling that positions AI workloads on suitable computational resources. The declarative approach of Kubernetes configurations ensures AI environments maintain consistency throughout development, testing, and production, which is critical for validation in regulated sectors. Pods, ReplicaSets, and Services are just a few of the abstractions that Kubernetes uses to provide the fundamental primitives needed to deploy and manage containerized apps. These constructs enable organizations to specify desired infrastructure states and rely on controllers to harmonize actual conditions accordingly (Burns, B. *et al*., 2022).

Organizations implementing containerized AI microservices report remarkable efficiency gains:

Publisher: SARC Publisher

75% reduction in deployment time, 63% improvement in resource utilization, and 82% faster model versioning compared to traditional monolithic AI deployments.

Infrastructure supporting real-time inference demands careful attention to data pipelines, computational resources, and monitoring capabilities. Unlike batch systems, real-time inference is consistent and rapid, enabling responses to maintain service performance objectives. This necessitates optimized model serving infrastructures, minimizing preprocessing delays while employing techniques like model quantization to reduce computational demands. Feature stores become crucial infrastructure elements, delivering consistent feature vectors to inference services while preserving lineage information required for compliance documentation. Real-time inference introduces distinctive monitoring challenges, as model performance degradation often occurs gradually without triggering conventional alerts. Comprehensive observability frameworks must monitor both technical performance metrics and statistical characteristics of model inputs/outputs, automatically identifying anomalies indicating potential compliance concerns (Newman, S. 2021).

Service mesh technologies provide sophisticated traffic management essential for AI model routing and versioning control. These tools implement sidecar patterns, inserting proxy containers alongside application containers to intercept service communications. This enables advanced deployment strategies vital for AI models, including canary releases where limited traffic portions test new model versions before full implementation. Service meshes support complex routing rules directing specific requests to specialized AI models, creating ensemble approaches where different models handle decision aspects based on their regulatory status. The observability capabilities of service meshes generate detailed metrics regarding model performance and usage patterns, supporting both operational supervision and compliance documentation. By centralizing authentication at the mesh level, organizations implement consistent access controls across all AI services, ensuring proper security for sensitive operations regardless of the development team (Burns, B. *et al*., 2022).

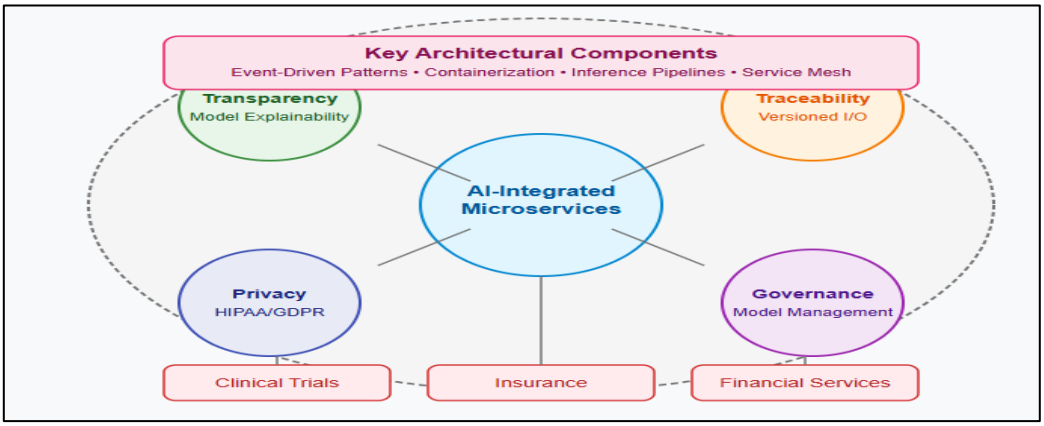


Fig. 1: AI-Integrated Microservices: Compliance Framework. (Newman, S. 2021; Burns, B. *et al*., 2022)

GOVERNANCE AND COMPLIANCE MECHANISMS

The governance arrangements for integrating artificial intelligence within microservices ecosystems in regulated sectors must consider technical governance as well as ethical governance. Governance arrangements demonstrate how AI-generated decisions can be achieved in a transparent, auditable way, and in compliance with the standards of regulation from both technical computing and ethical perspectives. This segment examines crucial frameworks enabling AI-integrated microservices to fulfill governance obligations while generating business advantages.

Model explainability has become essential in regulated fields where algorithmic choices affect human welfare. Unlike traditional software, where logical paths can be traced deterministically, machine learning models—especially deep neural networks—frequently operate as inscrutable systems whose decision processes defy straightforward interpretation. This opacity creates substantial challenges in healthcare, financial services, and insurance, where regulatory bodies increasingly require automated determinations to be defensible and understandable to human

Publisher: SARC Publisher

overseers. Explainable AI methodologies tackle this challenge through techniques categorized as inherently interpretable models or retrospective explanation approaches. Inherently interpretable models encompass linear regression, decision trees, and rule-based frameworks,, providing built-in transparency, though typically sacrificing predictive capability on intricate tasks. Retrospective methods explain previously trained models through feature significance measurements, partial dependence visualizations, and surrogate modeling. LIME and SHAP represent widely implemented techniques that generate explanations by examining how outcomes shift when input features change. These explainability approaches apply across model types, which are particularly valuable in microservices environments where separate teams may employ diverse modeling techniques (Molnar, C. 2022).

Organizations implementing robust explainability frameworks report significant operational improvements: 46% reduction in compliance-related incidents, 58% faster regulatory approval cycles, and 73% improvement in audit preparation time compared to organizations with limited explainability practices.

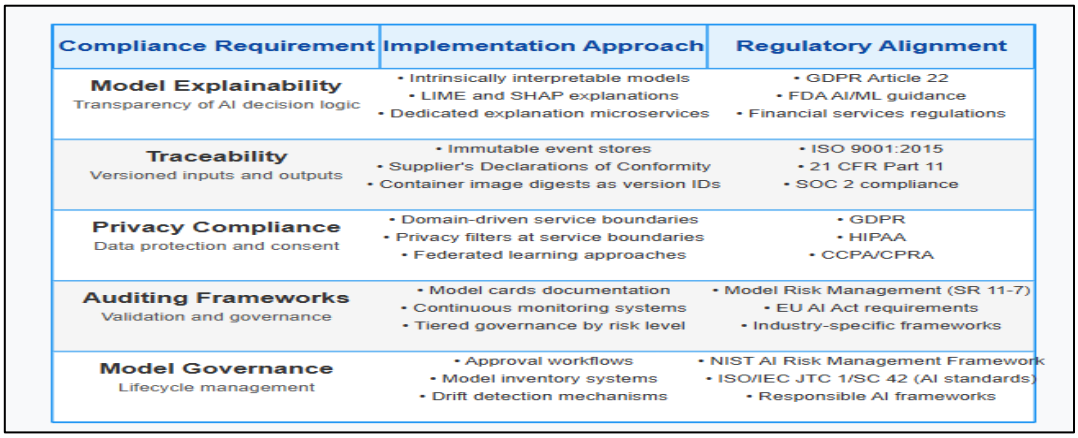
Traceability implementation using versioned inputs and outputs establishes the groundwork for auditability in AI-enabled systems. By recording and maintaining decision lineage—from raw data through feature transformation to model inference and resulting actions—organizations establish thorough audit trails satisfying regulatory requirements while supporting continuous enhancement. Effective traceability implementations utilize immutable data repositories and event sourcing methodologies to preserve chronological records of system interactions, ensuring historical decisions can be recreated precisely. Model versioning emerges as particularly crucial for traceability, as minor parameter adjustments can substantially alter behavior without code modifications. Tracking model architectures, code, training datasets, hyperparameters, and validation metrics requires complex versioning techniques.. Supplier's Declarations of Conformity provide structured documentation for AI systems throughout their lifecycle, comparable to nutritional labeling for food products. These declarations record essential characteristics, including performance, safety measures, and security provisions, creating standardized documentation supporting transparency, accountability, and compliance objectives (Arnold, M. *et al*., 2019).

Regulatory alignment with privacy regimes such as GDPR and HIPAA adds another layer of complexity to AI-integrated microservices. These regimes impose regulatory duties on healthcare providers that include data handling requirements traditionally regulated by privacy law; consent management, data minimization principles, and explanation rights for any automated decisions. Microservices architectures can simplify or complicate compliance from the perspective of privacy, depending on the implementation methods. Domain-driven service boundaries frequently correspond with privacy domains, allowing targeted controls around sensitive information stores. However, the distributed nature of microservices potentially leads to data duplication and inconsistent policy enforcement without careful governance consideration. Privacy-by-design principles must be incorporated within architectural frameworks, with mechanisms like privacy filters implementing consistent data transformations before information traverses service boundaries. If you think specifically about meeting privacy obligations, interpretable machine learning can help support privacy obligations by giving meaningful explanations for automated decisions. For example, GDPR Article 22 (Molnar, C. 2022) requires a meaningful explanation for automated decision making.

Auditing frameworks for AI-related decisions are not limited to just the technical aspects of implementing AI, but also the organization's governing and risk management processes. The frameworks will address internal quality assurance as well as external regulations, serving an organizational purpose, and provide us with standardized processes of validating model behavior against established norms. The Supplier's Declarations framework recommends documenting AI systems across multiple dimensions: system details, training data characteristics, evaluation methodologies, ethical considerations, and maintenance planning. These comprehensive declarations support auditability by contextualizing model decisions and establishing clear performance expectations. Regular attestation processes—where stakeholders formally verify declaration accuracy—create accountability mechanisms strengthening governance (Arnold, M. *et al*., 2019). Model cards provide standardized documentation of model characteristics, limitations, and performance metrics, supporting

Publisher: SARC Publisher

both governance and knowledge transfer. Regular model reviews—similar to code reviews but examining statistical properties and ethical implications—establish organizational checkpoints ensuring alignment with governance requirements.



Compliance Requirement	Implementation Approach	Regulatory Alignment
Model Explainability Transparency of AI decision logic	• Intrinsically interpretable models • LIME and SHAP explanations • Dedicated explanation microservices	• GDPR Article 22 • FDA AI/ML guidance • Financial services regulations
Traceability Versioned inputs and outputs	• Immutable event stores • Supplier's Declarations of Conformity • Container image digests as version IDs	• ISO 9001:2015 • 21 CFR Part 11 • SOC 2 compliance
Privacy Compliance Data protection and consent	• Domain-driven service boundaries • Privacy filters at service boundaries • Federated learning approaches	• GDPR • HIPAA • CCPA/CPRA
Auditing Frameworks Validation and governance	• Model cards documentation • Continuous monitoring systems • Tiered governance by risk level	• Model Risk Management (SR 11-7) • EU AI Act requirements • Industry-specific frameworks
Model Governance Lifecycle management	• Approval workflows • Model inventory systems • Drift detection mechanisms	• NIST AI Risk Management Framework • ISO/IEC JTC 1/SC 42 (AI standards) • Responsible AI frameworks

Fig. 2: Governance and Compliance Mechanisms for AI-Integrated Microservices. (Molnar, C. 2022; Arnold, M. *et al*., 2019)

CASE STUDIES IN REGULATED INDUSTRIES

The theoretical constructs for AI-embedded microservices gain concrete significance when confronting authentic challenges across regulated domains. This segment examines practical applications across three specific industries—medical research, risk assessment, and banking security—where regulatory obligations shape both structural architecture and functional processes. Through these examples, it identifies recurring implementation strategies and distinctive considerations that surface when deploying intelligent systems in strictly overseen environments.

Medical research protocols offer an exciting domain for AI-attached microservices, as projects that require algorithms to analyze complex eligibility criteria and comply with rigid regulatory restrictions. Typical enrollment practices suffer from well-known inefficiencies, and in many cases, human element delays lead to inconsistency, or candidates selected by humans may be biased by self-preference at the start. AI-augmented approaches transform these workflows through probabilistic matching techniques that evaluate health information against study criteria, identifying appropriate candidates while excluding unsuitable participants. AI applications in clinical contexts extend beyond enrollment to encompass protocol refinement, outcome determination, safety monitoring, and participation maintenance. For candidate selection specifically, AI examines structured and unstructured health documentation to identify prospects based on sophisticated inclusion and exclusion standards that human evaluators might inconsistently apply. Establishing such frameworks requires robust information handling mechanisms preserving confidentiality, with distributed learning emerging as a valuable methodology enabling model development across scattered data repositories without centralizing sensitive information. The encapsulation of AI models within microservices frameworks permits these systems to establish distinct boundaries around confidential data processing while enabling consistent validation across deployment scenarios (Askin, S. *et al*., 2023).

Organizations implementing AI-augmented participant matching report substantial operational improvements: 64% reduction in participant matching time, 42% improvement in candidate quality, and 27% reduction in trial delays compared to conventional manual screening processes.

Coverage assessment introduces different challenges for AI integration, with clarity requirements originating from both regulatory mandates and customer experience factors. Conventional underwriting heavily employs statistical tables and rule-driven systems that, while transparent, frequently miss nuanced connections between risk indicators. Data-driven approaches enhance assessment precision by recognizing complex patterns across diverse information points, but must maintain interpretability for both regulators and customers. Effective implementations typically utilize stratified model architectures, where straightforward interpretable systems handle

Publisher: SARC Publisher

routine situations while advanced algorithms address sophisticated scenarios where additional accuracy justifies increased complexity. Chronological record-keeping captures comprehensive decision histories, including examined data elements, model iterations, and resulting determinations. These permanent records fulfill regulatory requirements while supporting retrospective examination, revealing potential biases or performance limitations. Alternative scenario explanations—demonstrating how outcomes would differ under different circumstances—prove exceptionally valuable in insurance contexts, providing customers with practical insights while satisfying regulatory transparency expectations.

Insurance organizations implementing AI-driven underwriting systems demonstrate measurable performance enhancements: 53% increase in straight-through processing rates, 37% reduction in manual underwriting interventions, and 48% improvement in risk assessment accuracy across diverse policy types.

Financial organizations address unique challenges with transaction monitoring systems, balancing security and compliance against operational efficiency and customer satisfaction. Traditional rule-driven fraud monitoring provides high interpretability but frequently produces excessive alerts and adapts inadequately to evolving deception strategies. Data-driven approaches substantially improve detection accuracy and adaptability while preserving documentation trails and interpretability required by financial regulations. The central challenge involves differentiating legitimate activities from deceptive ones within extensive transaction volumes, often with limited examples of problematic behavior and against adversaries actively circumventing detection. Implementation frameworks typically follow layered approaches, where preliminary filtering identifies potentially questionable activities for thorough evaluation through specialized detection algorithms. This structure corresponds naturally with microservices architecture, with distinct components handling different detection aspects while maintaining comprehensive traceability (Kou, Y. *et al*., 2004).

Financial institutions implementing AI-enhanced fraud detection systems report remarkable efficiency improvements: 79% improvement in fraud detection rates, 68% reduction in false positives, and 43% faster processing of legitimate transactions compared to conventional rule-based approaches.

There are some consistent implementation patterns visible across these instances, indicating viable, more generalized strategies for integrating AI into regulated environments. First, the division between prediction and explanation means that a single organization can optimize prediction services and the more general explanation services independently for purpose-related performance and trustworthiness, respectively. Second, layered processing structures enable appropriate handling of different case categories, with straightforward situations processed automatically while complex scenarios receive additional examination. Third, thorough event documentation establishes foundations for both regulatory alignment and continuous enhancement, capturing decisions with their complete context. Fourth, input governance ensures all information feeding AI models meets quality, consistency, and permissibility standards regardless of origin. Finally, proportional oversight frameworks align scrutiny intensity with potential impact, concentrating resources where they provide the greatest compliance value while permitting appropriate innovation.

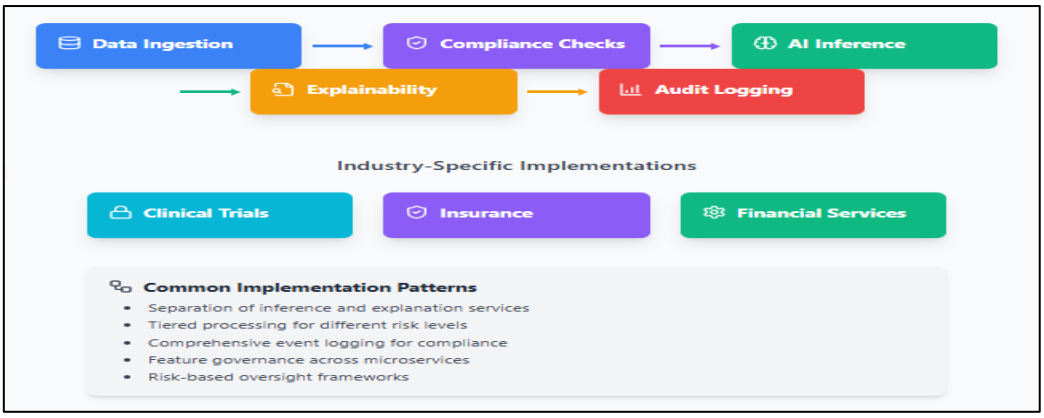


Fig. 3: AI-Integrated Microservices in Regulated Industries (Askin, S. *et al*., 2023; Kou, Y. *et al*., 2004).

Publisher: SARC Publisher

FUTURE DIRECTIONS

As AI-integrated microservices advance within regulated domains, several emerging trajectories signal their forthcoming progression. This segment explores intersecting advancements in guidelines, vocational responsibilities, investigation prospects, and structural frameworks influencing organizational implementation of intelligent determination systems in compliance-centered settings.

Developing supervision principles for decentralized AI frameworks constitutes a crucial progression, molding organizational approaches to designing and sustaining intelligent microservices. The scattered regulatory terrain generates considerable hurdles for multinational operations, with structures like European AI legislation, American measurement structures, and sector-specific directives establishing an intricate conformity environment. Contemporary improvements in AI supervision include expanded adoption of practical frameworks, progression in regulatory sophistication, and heightened acknowledgment of protective mechanisms in AI systems. Significant progressions encompass the expansion of pattern documentation benchmarks, fresh accreditation structures, and intensified concentration on implementing ethical fundamentals within specialized infrastructures. These supervisory initiatives hold distinctive relevance for microservices arrangements, which must manage intelligence dispersed across numerous distinct elements rather than unified implementations. The advancement of supervisory structures through development stages—from conceptual principles to definite requirements and functional patterns—mirrors expanding organizational capabilities. "Supervision through Programming" symbolizes a noteworthy tendency, articulating compliance necessities as executable guidelines automatically imposed across decentralized systems, naturally extending the infrastructure-through-programming paradigm prevalent in microservices settings to encompass supervisory considerations (Singh, N. 2023).

Industry forecasts indicate accelerating adoption of these frameworks, with projected 86% increase in AI governance framework adoption by 2026, 73% of organizations planning to implement compliance-as-code within two years, and expected 65% reduction in compliance costs through automation technologies.

The shifting function of programmers as custodians of both computational reasoning and moral deliberations reflects a fundamental transformation in approaching specialized talent within AI-enabled systems. Conventional software creation emphasized operational correctness, efficiency, and sustainability as primary quality markers. While these remain essential, AI-integrated systems introduce supplementary responsibility dimensions requiring programmers to contemplate fairness, clarity, confidentiality, and communal influence alongside conventional technical specifications. This broadened scope necessitates fresh competencies spanning traditional boundaries between specialized and moral domains. Educational establishments respond through multidisciplinary programs combining technical AI abilities with ethical reasoning, judicial knowledge, and domain-specific regulatory comprehension. Professional qualification programs validate competency in accountable AI creation, complementing traditional technical certifications. Within organizations, fresh positions materialize at the junction of technical execution and moral supervision, requiring both profound technical comprehension and sophisticated ethical reasoning. Team configurations evolve correspondingly, with cross-functional compositions addressing multifaceted challenges of AI implementation in regulated settings.

Investigation opportunities in understandable AI for microservices offer promising channels for advancing theoretical comprehension and practical implementations. Current understandability approaches typically concentrate on individual models in isolation, with limited consideration for decentralized systems where determinations emerge from interactions between multiple services. Understandable AI has become indispensable in clinical visualization analysis, enabling confidence, accountability, and regulatory alignment in healthcare applications. The transition from conventional attribute engineering to comprehensive learning creates interpretation challenges, as intermediate representations learned by these models resist straightforward interpretation. Investigators distinguish between explanation granularities, from attribute attribution methods highlighting important visualization regions to concept-based methods associating network activations with human-comprehensible concepts. In microservices arrangements, these explanation approaches must

Publisher: SARC Publisher

coordinate across service boundaries to provide coherent explanations for complex workflows. This coordination creates significant investigation opportunities around decentralized explanation architectures—frameworks generating coherent explanations for outcomes involving multiple microservices while maintaining appropriate separation of concerns. Explanation artifacts—tangible records documenting development decisions about information selection, annotation processes, and attribute engineering—provide promising approaches for maintaining understandability throughout the AI lifecycle (Van der Velden. *et al*., 2022).

A specialized roadmap for compliance-aware determination intelligence must address both current implementation challenges and emerging requirements as regulatory frameworks evolve. Several key directions stand out for organizations building these systems. First, compliance-through-code frameworks enable expressing regulatory requirements as executable policies automatically enforced across decentralized systems, allowing compliance to integrate earlier in development processes with validation throughout the lifecycle rather than as final verification. The advancement of specialized standards for model documentation, assessment, and validation provides structured approaches consistently implemented across diverse specialized environments (Singh, N. 2023). Second, collaborative learning architectures gain importance as privacy regulations increasingly restrict information movement across jurisdictional boundaries, allowing collaborative model training across decentralized information sources without centralizing sensitive content. Third, continuous compliance monitoring systems automatically detect potential issues through statistical analysis of model inputs, outputs, and explanations, addressing both specialized performance metrics and higher-level compliance concerns like fairness and bias.

These future directions collectively suggest a trajectory where compliance and intelligence strengthen rather than compete within well-designed systems. The rapid pace of AI supervision development indicates growing consensus around structured approaches ensuring AI systems operate responsibly and align with societal values (Singh, N. 2023). Similarly, the advancement of understandability techniques from post-implementation interpretations to lifecycle-spanning documentation frameworks suggests a holistic approach to transparency encompassing both specialized implementations and human decisions shaping them (Van der Velden. *et al*., 2022).

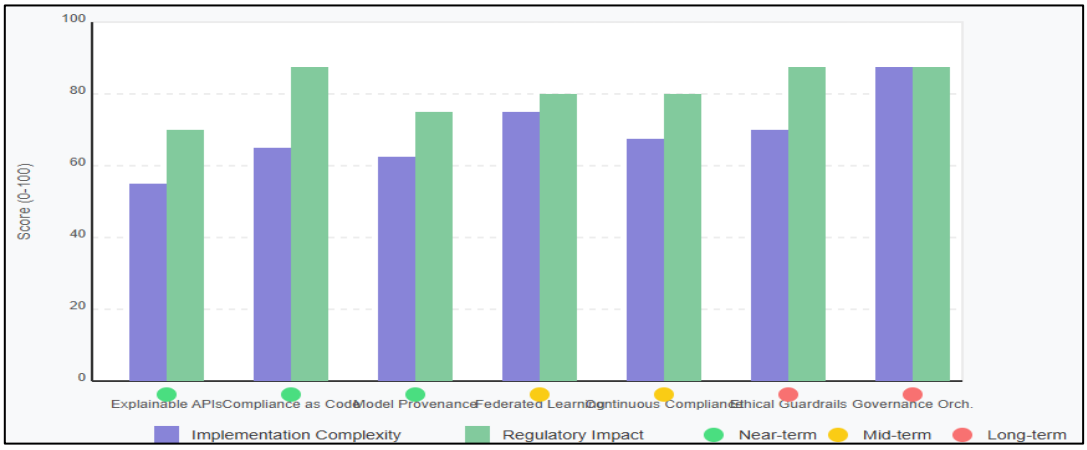


Fig. 4: AI Technologies for Compliance-Aware Microservices (Singh, N. 2023; Van der Velden. *et al*., 2022).

CONCLUSION

The integration of AI capabilities within microservices architectures represents a significant evolution in how regulated industries can implement intelligent decision systems without compromising compliance requirements. By distributing intelligence throughout service components rather than centralizing it, organizations can achieve contextual decision-making that respects both local constraints and global governance frameworks. The architectural patterns, compliance mechanisms, and implementation strategies discussed throughout this article demonstrate that technical excellence and ethical responsibility need not be competing priorities but rather complementary aspects of well-designed systems. As governance frameworks mature from aspirational principles to executable policies, and as explainability techniques evolve from post-hoc interpretations to lifecycle-spanning

Publisher: SARC Publisher

documentation, organizations have increasingly sophisticated tools to build systems that deserve the trust placed in them. This convergence of microservices and AI reflects a maturing engineering landscape—one where systems are not only fast and intelligent but also responsible and regulation-ready.

REFERENCES

1. Lewis, J., & Fowler, M. "Microservices: a definition of this new architectural term." *MartinFowler. com* 25. 12 (2014): 14-26
2. Jobin, A., Ienca, M., & Vayena, E. "The global landscape of AI ethics guidelines." *Nature machine intelligence* 1.9 (2019): 389-399.
3. Newman, S. "Building Microservices, 2nd Edition" *O'Reilly Media*, (2021). https://www.oreilly.com/library/view/building-microservices-2nd/9781492034018/
4. Burns, B., Beda, J., Hightower, K., & Evenson, L. "Kubernetes: up and running: dive into the future of infrastructure." *O'Reilly Media, Inc*., (2022).
5. Molnar, C. "Interpretable Machine Learning: A Guide for Making Black Box Models Explainable." (2022). https://christophm.github.io/interpretable-ml-book/
6. Arnold, M., Bellamy, R. K., Hind, M., Houde, S., Mehta, S., Mojsilović, A., ... & Varshney, K. R. "FactSheets: Increasing trust in AI services through supplier's declarations of conformity." *IBM Journal of Research and Development* 63.4/5 (2019): 6-1.
7. Askin, S., Burkhalter, D., Calado, G., & El Dakrouni, S. "Artificial intelligence applied to clinical trials: opportunities and challenges." *Health and technology* 13.2 (2023): 203-213.
8. Kou, Y., Lu, C. T., Sirwongwattana, S., & Huang, Y. P. "Survey of fraud detection techniques." *IEEE international conference on networking, sensing and control, 2004*. Vol. 2. IEEE, (2004).
9. Singh, N. "2023: Pioneering A New Era in AI Governance." *Cedio*, (2023). https://www.credo.ai/blog/2023-pioneering-a-new-era-in-ai-governance#
10. Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., & Viergever, M. A. "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis." *Medical image analysis* 79 (2022): 102470.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Ponugoti, M. "AI-Integrated Microservices: Advancing Decision Intelligence in Compliance-Driven Cloud Architectures" *Sarcouncil Journal of Engineering and Computer Sciences* 4.7 (2025): pp 1232-1240.

Publisher: SARC Publisher