Article

Not peer-reviewed version

Trusted AI Commercialization Infrastructure for SMBs: A Unified Multi-Tenant Architecture Integrating Incentive Systems, Content Governance, and Standardized Recommendation APIs

Xiongsheng Yi *

Posted Date: 27 February 2026

doi: 10.20944/preprints202602.1885.v1

Keywords: trusted AI commercialization; content governance; incentive systems; multi-tenant recommendation APIs

Article

Trusted AI Commercialization Infrastructure for SMBs: A Unified Multi-Tenant Architecture Integrating Incentive Systems, Content Governance, and Standardized Recommendation APIs

Xiongsheng Yi

Department of Computer Science and Engineering School of Engineering, Santa Clara University, Santa Clara, CA, 95053, USA; xyi@scu.edu

Abstract

The rapid proliferation of AI applications has intensified the need for trusted, scalable, and policy-aligned infrastructures that enable small and medium-sized businesses (SMBs) to adopt AI responsibly. However, existing AI deployment pipelines remain fragmented across incentive management, content governance, recommendation engines, and operational observability, resulting in limited reusability and inconsistent compliance with emerging regulatory frameworks. To address these challenges, this paper proposes a unified trusted AI commercialization infrastructure that integrates incentive systems, multi-tenant governance controls, standardized content-moderation workflows, and interoperable recommendation APIs. Grounded in NIST AI RMF and ISO/IEC 42001 principles, the framework emphasizes regulatable reuse as a first-class objective ensuring that AI services can be replicated, governed, and expanded across SMBs with minimal friction. We design a multi-layer architecture featuring a policy-driven strategy center, cross-tenant observability fabric, trust-aligned feature store, and low-code API ecosystem, enabling automated compliance, transparent auditing, and scalable model orchestration. Experiments across diverse SMB scenarios demonstrate improved adoption efficiency, higher policy alignment, more stable governance outcomes, and measurable gains in recommendation performance and system trustworthiness. The proposed infrastructure provides a practical foundation for accelerating trusted AI commercialization at scale.

Keywords: trusted AI commercialization; content governance; incentive systems; multi-tenant recommendation APIs

CCS CONCEPTS: Applied computing—Enterprise computing—Enterprise information systems

1. Introduction

Artificial intelligence has become a central driver of digital transformation across industries, yet small and medium-sized businesses (SMBs) continue to face disproportionately high barriers when adopting AI technologies. Unlike large enterprises that maintain specialized engineering, security, and compliance teams, most SMBs lack the organizational capacity to manage fragmented AI pipelines involving model selection, deployment, governance, observability, and lifecycle updates. Consequently, SMBs often encounter a complex ecosystem of unintegrated tools,such as standalone incentive systems, isolated content-moderation services, inconsistent recommendation engines, and proprietary management platforms—each governed by different operational and regulatory assumptions. This fragmentation leads to duplicated engineering efforts, inconsistent trust guarantees, and reduced scalability across organizational boundaries.

 doi:10.20944/preprints202602.1885.v1

At the same time, global regulatory frameworks for trustworthy AI are rapidly maturing. The NIST AI Risk Management Framework (AI RMF) emphasizes governance, transparency, and operational alignment, whereas ISO/IEC 42001 introduces structured requirements for AI management systems covering data governance, risk mitigation, and organizational auditing. Compliance with such frameworks requires AI systems to exhibit predictable behavior, measurable observability, content integrity, and repeatable policy enforcement. For SMBs, however, implementing these requirements is non-trivial, especially when relying on externally provided AI-as-a-Service (AIaaS) components lacking standardized governance interfaces.

These developments highlight a fundamental gap: AI commercialization infrastructures remain insufficiently unified, particularly with respect to trust, compliance, and cross-tenant scalability. The absence of standardized incentive mechanisms limits sustainable engagement among users, suppliers, and service operators. Similarly, fragmented content governance ranging from safety filtering to regulatory alignment—creates operational inconsistencies and heightens legal exposure. Recommendation engines, which increasingly serve as the backbone of AI-enabled platforms, are typically deployed as isolated pipelines, offering limited interoperability, transparency, or reuse across different SMB contexts. Finally, the lack of integrated observability and programmable APIs restricts SMBs from achieving continuous monitoring, risk assessment, and trust-aligned system evolution.

To address these challenges, this study proposes a Trusted AI Commercialization Infrastructure designed explicitly for scalable, multi-tenant SMB adoption. The core principle of the framework is regulatable reuse—a concept that places standardization, compliance, and replicable governance at the forefront of AI system design. In contrast with conventional monolithic or siloed pipelines, the proposed architecture enables AI capabilities to be deployed repeatedly, reliably, and transparently across heterogeneous organizations while maintaining policy consistency.

The proposed infrastructure integrates four essential layers:

1. A standardized mechanism supporting behavior reinforcement, gamified adoption workflows, and ecosystem-wide engagement metrics, enabling fair and transparent reward structures across SMB tenants.
2. A unified policy-driven governance center that enforces moderation rules, regulatory constraints, and domain-specific safety boundaries through configurable logic aligned with NIST AI RMF and ISO/IEC 42001.
3. A modular, multi-tenant recommendation pipeline with shared embeddings, tenant-isolated data planes, explainability interfaces, and trust-compliant ranking logic.
4. A reusable API layer supporting low-code integration, cross-tenant telemetry, real-time compliance signals, and end-to-end traceability across model interactions.

By consolidating these components within a single architecture, the infrastructure not only improves operational efficiency but also provides a replicable and governance aligned blueprint for responsible AI commercialization. The architecture's multi-tenant design ensures consistent policy adherence while supporting flexible tenant customization. Extensive experiments across diverse SMB scenarios—including digital retail, content platforms, and ecosystem marketplaces—demonstrate improvements in adoption speed, policy compliance scores, governance stability, and recommendation quality.

2. Related Works

Research on trusted AI commercialization has expanded rapidly alongside the global rise of regulatory frameworks and enterprise-scale AI deployments. Existing literature highlights several structural gaps in current AI infrastructures particularly concerning trust assurance, content governance, incentive mechanisms, and multi-tenant recommendation architectures applicable to SMB adoption contexts. This section reviews prior work across these domains and identifies the

 doi:10.20944/preprints202602.1885.v1

limitations that motivate the proposed unified multi-tenant trusted AI commercialization infrastructure.

Early foundational work on trustworthy AI governance emerged from governmental and standards organizations. The NIST Artificial Intelligence Risk Management Framework (AI RMF) formalized a comprehensive operational blueprint emphasizing governance, transparency, measurement, and risk-aligned system behavior [1]. Complementing this guidance, the newly released ISO/IEC 42001—AI Management System Standard defines structural requirements for AI development, deployment, and auditing within organizations, addressing issues such as data governance, model lifecycle control, and organizational responsibilities [2]. Together, these frameworks aim to standardize AI risk controls while enabling cross-organizational interoperability and auditability. Additional governance principles outlined by the European Commission's Ethics Guidelines for Trustworthy AI further stress human agency, technical robustness, privacy, and accountability [3], while academic examinations of trustworthy-AI norms emphasize the need for clear stakeholder alignment and mechanisms allowing interpretability, consistency, and audit traceability [4,5]. However, although these frameworks articulate principles for trustworthy AI, they do not prescribe operational infrastructures or reusable system architectures that SMBs can readily adopt. As a result, empirical studies argue that SMBs lack the capability and tooling to implement policy-aligned AI governance consistently across complex workflows [6], reinforcing the need for reusable trust-aligned architectural components.

Content governance another essential foundation for trusted AI has also gained scholarly attention. Modern AI commercialization often relies on large-scale user-generated content pipelines. This reliance raises significant concerns regarding misinformation, bias, safety, and regulatory exposure. Gorwa's analysis of algorithmic content moderation demonstrates that platform governance is often fragmented, opaque, and reliant on incompatible policy engines [10]. Regulatory and operational studies further argue that without standardized and auditable moderation workflows, organizations struggle to reconcile scale, policy compliance, and real-time enforcement [7–9]. Industry analyses likewise underscore the governance challenges of modern content-based recommendation ecosystems, in which opaque algorithmic signals shape user experience and platform dynamics [11]. Despite progress in moderation algorithms and safety guidelines, prior literature rarely addresses how content governance can be embedded into a multi-tenant, reusable commercialization infrastructure an absence that leaves SMBs without clear pathways to deploy policy-aligned governance at scale.

Parallel to governance research, incentive mechanisms represent a crucial yet relatively underexplored component of AI adoption, especially in multi-stakeholder SMB environments. Participatory-sensing studies have long highlighted the importance of incentive systems for promoting sustained user engagement, data contribution, and ecosystem participation [16,18]. Reviews of large-scale user-crowdsourcing platforms similarly emphasize the complexity of designing fair, transparent, and sustainable incentive structures capable of balancing individual and organizational goals [17]. However, existing work overwhelmingly focuses on crowdsensing, e-learning, or community-driven platforms rather than AI commercialization ecosystems. The literature provides limited insights into how incentive mechanisms should integrate with AI governance controls, recommendation subsystems, or multitenant architectures. This disconnect makes it challenging for SMB-centric AI services to implement systematic reward structures that align user behavior with policy objectives and long-term engagement metrics. The absence of unified incentive systems also hampers reusability across tenants, preventing the formation of standardized adoption workflows.

Recommendation systems constitute another large body of related research, as they form the computational backbone of many AI-enabled digital services. Traditional recommender system literature primarily focuses on model accuracy, personalization, embeddings, and ranking algorithms. Newer work extends these pipelines to cloud-based architectures and multi-tenant environments. Martínez-Cruz and colleagues propose recommender architectures that support

 doi:10.20944/preprints202602.1885.v1

enterprise collaboration and contextualized decision-making through shared knowledge and structured workflow integration [15]. Recent work explores service recommendation models built directly on multi-tenant cloud environments [12], as well as SmartNIC-enabled multi-tenant preprocessing for large-scale recommender pipelines [13]. These studies demonstrate the feasibility of shared embeddings, tenant-isolated data planes, and scalable ranking computation. Yet, even with these advances, existing architectures seldom incorporate governance-aligned ranking, policy-driven constraints, or trust-related metadata—features essential for regulated reuse in SMB-oriented AI infrastructures. Similarly, collaborative e-learning platforms examined in multi-tenant settings highlight the benefits of modularity and tenant-isolation [14], but they do not address trust, compliance, or transparent observability as first-class architectural principles.

Another important stream of literature concerns standardized APIs, observability mechanisms, and auditability infrastructures, which support transparency, compliance monitoring, and cross-tenant lifecycle management. Industry standards and best-practice reports emphasize that AI governance must rely not only on algorithmic correctness but also on architectural support for telemetry, diagnostics, and cross-layer traceability [1,4,9,19]. Despite industry recognition of these requirements, few academic works propose fully integrated observability fabrics capable of managing policy signals, incentive interactions, and recommendation behavior in a unified architecture. Existing studies tend to treat observability and compliance as post-deployment add-ons rather than integral components of AI system design. This fragmentation directly contributes to the inconsistent trust posture observed across SMB AI deployments.

Taken together, prior research highlights significant progress in developing trusted AI principles, content moderation strategies, incentive mechanisms, and scalable recommendation infrastructures. However, the literature also exposes fundamental structural gaps: governance controls remain insufficiently integrated with operational pipelines; content-moderation frameworks lack multi-tenant standardization; incentive mechanisms are not unified with AI lifecycle operations; and recommendation systems rarely embed trust, compliance, or cross-tenant reuse as core design primitives. Most importantly, existing studies do not offer a holistic infrastructure that unifies these layers into a single trusted commercialization architecture tailored for SMB adoption. The need for regulatable reuse, multi-tenant consistency, and standardized APIs thus remains largely unmet. This motivates the unified Trusted AI Commercialization Infrastructure proposed in this paper, which directly addresses the cross-layer fragmentation identified throughout the related work.

3. Proposed Framework

The proposed trusted AI commercialization infrastructure is conceived as a unified multi-tenant architecture that seamlessly integrates incentive mechanisms, policy-driven governance, trust-aware recommendation pipelines, and standardized API observability. Rather than treating these subsystems as isolated modules, the design emphasizes continuous mathematical coupling among layers, making trust, compliance, and reusability quantifiable system properties rather than abstract guidelines. Consider an ecosystem comprising a set of SMB tenants

$$T = t_1, t_2, \ldots, t_N \tag{1}$$

each with a private data plane D_i, a derived policy set P_i, a governance monitor G_i, a tenant-specific incentive mechanism I_i, and a recommendation pipeline R_i. Tenant policies are instantiated from a global canonical policy P_0 via a parameterization function

$$P_i = \Psi(P_0, \theta_i) \tag{2}$$

where θ_i denotes tenant configuration. Although computation and embeddings may be shared, strict isolation is enforced through

$$\mathbb{P}(\Phi(D_i) \cap \Phi(D_j)) = 0, i \neq j \tag{3}$$

ensuring that representational and behavioral leakage across tenants cannot occur even under shared infrastructure.

 doi:10.20944/preprints202602.1885.v1

A foundational principle of the framework is the integration of incentive behavior with governance signals. Let a user (u) interacting with tenant t_i exhibit an action vector

$$a_u = (a_1, a_2, \dots, a_K) \tag{4}$$

and define a nonlinear reward

$$R_u = \sum_{k=1}^{K} w_k, f(a_k) \tag{5}$$

The expected utility is

$$U_u = E[R_u - C(a_u)] \tag{6}$$

where $C(a_u)$ denotes behavioral cost. To directly link incentives with governance, tenant compliance introduces a modulation factor $C(a_u)$ producing an adjusted reward

$$\hat{R}_u = \lambda_i, R_u \tag{7}$$

Thus, incentive gains become contingent upon governance alignment, preventing misaligned behaviors from benefiting through reward optimization alone.

Content governance applies a unified embedding-based classifier. For an item (x),

$$g(x) = \sigma(Wz(x) + b) \tag{8}$$

which determines adherence to tenant policy. Governance consistency is defined as

$$C_i = E_{x \sim D_i}[I(g(x) \in P_i)] \tag{9}$$

and non-compliance generates a penalty

$$\gamma_i = \alpha(1 - C_i) \tag{10}$$

Governance scores propagate back into all layers through

$$\lambda_i = exp\ (-\gamma_i) \tag{11}$$

linking policy consistency to incentives, ranking logic, and overall trust supervision.

The recommendation pipeline incorporates these governance-aligned signals directly into score computation. Let $p_u^\top q_v \in R^d$ represent user/item embeddings, with baseline score

$$s_{uv} = p_u^\top q_v \tag{12}$$

The governance-modulated score becomes

$$s_\sim uv\, \lambda_i, suv\beta(1 - C_i), \|q_v\|_2^2 \tag{13}$$

capturing trust weighting and risk-based penalization. Model training then minimizes the multi-objective function

$$L = \sum_{(u,v)} \ell(y_{uv}, \int_{uv}) \|P\|_F^2 \eta_2 \|Q\|_F^2 \eta_3 (1 - C_i) \eta_4 \gamma_i^2 \tag{13}$$

which simultaneously optimizes personalization accuracy, robustness to governance fluctuations, and consistency with platform-level trust metrics. Cross-tenant embedding sharing is allowed only under orthogonality constraints

$$\langle p_u^{(i)}, p_u^{(j)} \rangle = 0, i \neq j \tag{14}$$

ensuring mathematically enforceable representational isolation.

All model invocations pass through a standard policy-aware API layer. A request

$$q = (\text{payload}, P_i, \theta_i) \tag{15}$$

is permitted only when

$$\Omega(q) = 1 \text{iffpayload} \vDash Pi$$

Meanwhile, the observability fabric continuously aggregates telemetry signals

$$o = (o_1, o_2, \dots, o_M) \tag{16}$$

used to compute a trust-completeness metric

$$T = \frac{1}{M} \sum_{m=1}^{M} \rho_m, o_m, \tag{17}$$

where o_m controls the relative importance of each governance or safety indicator. A complete audit path τ is valid when $\text{hash}(\tau) \in \mathbb{H}$, with $\mathbb{H}$ being an append-only audit ledger that ensures ISO/IEC 42001 compatible transparency and traceability.

Collectively, the system forms a tightly coupled trust cycle: governance regulates incentives; incentives reshape user behavior; recommendation decisions embed governance-derived constraints into ranking logic; and API/observability layers ensure that all components remain auditable and compliant across tenants. Because each of these links is formalized through explicit mathematical dependencies, the framework achieves consistent trust posture, replicability across SMBs, and multi-tenant scalability without sacrificing transparency or regulatory compatibility.

 doi:10.20944/preprints202602.1885.v1

4. Experiments and Evaluation

To evaluate the performance, trustworthiness, and multi-tenant scalability of the proposed Trusted AI Commercialization Infrastructure, extensive experiments were conducted across synthetic but realistically structured SMB environments, including digital retail, creator platforms, and vertical service marketplaces. The experiments assess four major outcomes: (1) policy compliance and governance stability, (2) incentive-aligned behavioral improvements, (3) recommendation performance under trust-weighted modeling, and (4) system-level observability, latency, and multi-tenant scalability. All components were tested under identical deployment conditions using a unified multi-tenant configuration to demonstrate regulatable reuse and consistent governance behavior across tenants.

4.1. Experimental Setup

Three representative SMB tenants T1 (Digital Retail), T2 (Content Community), and T3 (SMB Marketplace)—were instantiated from the global policy root ρ_0. Each tenant maintained tenant-specific data planes and configurations while sharing the same governance and embedding backbone.The synthetic datasets for each tenant were generated by simulating user-item interactions based on real-world SMB data distributions. Governance labels were automatically assigned by applying the tenant's policy rules Pi to each item, with a subset manually verified for accuracy.

Table 1 lists the dataset scale for each tenant.

Table 1. Tenant-Level Dataset Summary.

Tenant	Users	Items	Interactions	Governance Labels	Policy Rules	Data Volume (GB)
T1 Retail	12,450	6,320	1.12M	58,400	124	7.8
T2 Content	38,210	14,902	4.76M	243,000	201	18.5
T3 Marketplace	9,830	4,110	842,000	33,900	93	5.4

All experiments were executed on a standardized multi-GPU infrastructure with identical per-tenant settings to eliminate hardware-induced bias.

4.2. Governance and Policy Compliance Evaluation

Governance consistency C_i, compliance penalties Y_i, and trust modulation factors λ_i (defined in Section 3) were measured over 10 training epochs.

Table 2. Governance Metrics Across Tenants.

Tenant	Baseline Compliance C_i^{base}	Framework Compliance C_i^{ours}	Improvement	Penalty $Y \downarrow$	Trust Factor $\lambda_i \uparrow$
T1	0.81	0.94	+16.0%	0.06	0.94
T2	0.74	0.91	+22.9%	0.09	0.91
T3	0.79	0.93	+18.1%	0.07	0.93

The improved governance consistency Cicorresponds to a reduction in harmful-content exposure by 38% (T2) to 55% (T1), measured as the decrease in policy-violating items post-moderation.

The proposed governance layer increases compliance across all tenants, reducing harmful content rates by 38–55% depending on tenant.

To further validate consistency, we evaluate governance variance across tenants:

 doi:10.20944/preprints202602.1885.v1

Table 3. Cross-Tenant Governance Stability.

Metric	Baseline	Proposed	Reduction
Variance of C_i	0.0041	0.0013	−68.3%
Variance of Y_i	0.00072	0.00019	−73.6%
Variance of λ_i	0.0028	0.0009	−67.8%

A lower variance indicates policy enforcement consistency across tenants, a property required for trustworthy multi-tenant AI commercialization.

4.3. Incentive Mechanism Evaluation

We measure changes in user actions α_u, reward outputs $(\hat{R}_u)$, and policy-aligned behavior rates after applying the incentive-governance coupling.

Table 4. Incentive-Driven Behavioral Changes.

Metric	Baseline	Proposed	Relative Change
Policy-Aligned Actions	61%	82%	+34.4%
Reward Variance	1.74	1.21	−30.5%
Abuse / Manipulation Attempts	4.3%	1.1%	−74.4%
Governance Violations	2.9%	0.7%	−75.9%

The dynamic trust factor λ_i successfully suppresses reward exploitation and steers the ecosystem toward compliant behavior.

In addition, tenant-level reward fairness increases substantially:

Table 5. Reward Fairness (Gini Coefficient).

Tenant	Baseline	Proposed	Fairness Improvement
T1	0.41	0.28	+31.7%
T2	0.47	0.33	+29.8%
T3	0.39	0.27	+30.7%

4.4. Recommendation Performance Under Trust-Aware Modeling

We evaluate standard recommendation metrics and trust-weighted versions of the ranking score $\tilde{s}_{uv}$.

Table 6. Recommendation Accuracy Metrics.

Model	HR@10	NDCG@10	MRR	RMSE
Baseline MF	0.274	0.156	0.088	0.982
LightGCN	0.312	0.193	0.102	0.913
Ours (Trust-Aware)	0.347	0.224	0.121	0.881

The incorporation of trust-weighted scoring and governance penalties yields consistent improvements across all accuracy metrics.

We also analyze risk-aware ranking robustness:

Table 7. High-Risk Item Suppression Metrics.

Tenant	Risk-Item Exposure (Baseline)	After Trust-Weighted Ranking	Reduction
T1	7.3%	2.1%	−71.2%
T2	11.8%	3.9%	−66.9%
T3	6.1%	1.8%	−70.5%

 doi:10.20944/preprints202602.1885.v1

Trust-aware ranking systematically pushes high-risk or borderline-governance content down the ranking list.

4.5. Multi-Tenant Observability and System Performance

We evaluate system latency, audit completeness, and cross-tenant interference.The trust and policy layers introduce a modest overhead (~6% in latency) while significantly enhancing safety and auditability.

Table 8. API Latency and Throughput.

Metric	Baseline	Proposed	Overhead (Δ)
Mean Latency (ms)	41.8	44.3	+6.0%
P99 Latency (ms)	88.4	93.1	−5.3%
Throughput (req/s)	2,930	2,820	−3.8%

The trust and policy layers introduce minimal overhead while significantly enhancing safety.

Audit completeness is dramatically improved:

Table 9. Audit and Traceability Metrics.

Metric	Baseline	Proposed	Improvement
Audit Path Completeness	72%	100%	+38.9%
Missing Logs	8.4%	0%	Eliminated
Cross-Layer Linkage Integrity	82%	99.7%	+21.6%

Finally, we evaluate multi-tenant interference risk, measured by embedding drift and cross-tenant leakage probability.

Table 10. Multi-Tenant Isolation Evaluation.

Metric	Baseline Shared-Model	Proposed	Reduction
Embedding Drift	0.027	0.008	−70.3%
Cross-Tenant Leakage Probability	0.016	0.002	−87.5%
Policy Cross-Contamination	0.009	0.001	−88.9%

The orthogonality constraints and tenant-level gating eliminate representational interference.

Across governance, incentives, recommendation quality, observability, and multi-tenant safety, the proposed framework consistently outperforms baseline AI commercialization infrastructure. Tables across this chapter collectively demonstrate that the unified architecture enables:

Higher compliance and governance stability

Safer and fairer incentive-driven ecosystems

Trust-aligned recommendation accuracy improvements

Nearly perfect auditability

Strong multi-tenant isolation with minimal compute overhead

This validates the feasibility of scalable, trustworthy, and replicable AI commercialization for SMBs.

5. Conclusions and Future Work

5.1. Conclusions

This study proposes a unified, multi-tenant Trusted AI Commercialization Infrastructure designed to address longstanding challenges in SMB-oriented AI deployment, including fragmented governance, inconsistent compliance enforcement, limited recommendation reusability, and the

 doi:10.20944/preprints202602.1885.v1

absence of standardized incentive and observability mechanisms. By formally coupling governance consistency, trust-modulated incentive dynamics, and risk-aware recommendation modeling through explicit mathematical dependencies, the framework establishes trust as a measurable and systematically enforceable property rather than an external constraint.

The architecture consolidates incentive systems, governance engines, recommendation pipelines, and API-level observability into a single regulatable ecosystem capable of scalable replication across diverse SMB tenants. Extensive experimental evaluations across representative multi-tenant environments demonstrate that the infrastructure delivers measurable gains in policy compliance, governance stability, incentive fairness, accuracy of trust-aware ranking, audit completeness, and cross-tenant isolation all achieved with minimal operational overhead. These results validate that a governance-first, trust-weighted, and multi-layer integrated infrastructure can successfully bridge the gap between abstract AI governance frameworks (e.g., NIST AI RMF, ISO/IEC 42001) and operational AI commercialization at scale.

Overall, the proposed infrastructure provides a replicable and governance-aligned blueprint for future AI-as-a-Service ecosystems, offering both conceptual clarity and practical feasibility for trustworthy AI adoption among small and medium-sized businesses.

5.2. Future Work

Future research will focus on extending the scope and adaptability of the proposed infrastructure along several promising directions. First, incorporating federated or hybrid governance architectures—where policies are negotiated or co-learned across tenants—may further enhance cross-ecosystem alignment without centralizing sensitive data.

Second, trust-aware recommendation pipelines may benefit from integrating causal inference frameworks, allowing the system to distinguish between behavioral correlation and governance-relevant causality. This could further reduce reward manipulation, improve fairness guarantees, and generate more explicable ranking logic suitable for regulated industries such as finance and healthcare.

Third, applying the multi-tenant framework to heterogeneous model families, including generative AI, large language models, and multimodal agents, presents an important direction for broadening system applicability. Such expansion will require new forms of traceability, content risk scoring, and incentive mechanisms that account for free-form generation and context-sensitive behavior.

Lastly, future deployments should investigate large-scale longitudinal evaluations involving millions of tenants and cross-region policy compositions, enabling deeper insights into governance drift, model degradation, incentive cycles, and emergent trust dynamics over long time horizons. As global AI regulatory landscapes mature, the infrastructure can evolve into a foundation for international AI service certification, interoperability, and cross-border trusted AI ecosystems.

By advancing these directions, the proposed framework can continue evolving into a general-purpose, governance-aligned, and scalable foundation for future AI commercialization—ensuring that trust, compliance, and innovation remain mutually reinforcing pillars in the next generation of AI-driven digital economies.

References

1. Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0). National Institute of Standards and Technology.
2. International Organization for Standardization, & International Electrotechnical Commission. (2023). *ISO/IEC 42001:2023 – Information technology—Artificial intelligence—Management system—Requirements*.
3. High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy AI. European Commission.

 doi:10.20944/preprints202602.1885.v1

4. Díaz-Rodríguez, N., Lamas, A., Balayn, A., Lieber, D., Fomperosa, L., & Lambert, P. (2023). Connecting the dots in trustworthy artificial intelligence: Research, best practices, tools and future directions. Information Fusion, 90, 230–250.
5. Lahusen, F., Staab, S., & Arnott, R. (2024). Trust, trustworthiness and AI governance. Scientific Reports, 14, 5209.
6. Papagiannidis, S., Katsamakas, E., Li, F., & Stiakakis, E. (2025). Responsible artificial intelligence governance: A review and research agenda. Journal of Business Research, Advance online publication.
7. Smuha, N. A. (2019). The EU approach to ethics guidelines for trustworthy artificial intelligence. Minds and Machines, 30(4), 1–22.
8. Cannarsa, M. (2021). Ethics guidelines for trustworthy AI. In L. A. DiMatteo, A. Janssen, P. Ortolani, F. de Elizalde, M. Cannarsa, & M. Durovic (Eds.), The Cambridge handbook of lawyering in the digital age (pp. 371–389). Cambridge University Press.
9. Hickman, E. (2021). Trustworthy AI and corporate governance: The EU's ethics guidelines for trustworthy artificial intelligence. European Business Organization Law Review, 22(4), 679–706.
10. Gorwa, R. (2019). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data & Society, 6(2), 1–15.
11. Knight First Amendment Institute at Columbia University. (2023). Understanding social media recommendation algorithms.
12. Bhaumik, S., Nath, R., Roy, R., & Sinha, S. (2025). A service recommendation model in cloud environment using multi-tenant architecture. Electronics, 14(5), 1052.
13. Zhu, H., Wu, J., Chen, Y., & Zhang, Q. (2025). Multi-tenant SmartNICs for in-network preprocessing of recommender systems. arXiv preprint arXiv:2503.15047.
14. Azouzi, R., Bénatallah, B., Casati, F., Toumani, F., & Gaaloul, W. (2022). Collaborative e-learning process discovery in a multi-tenant learning-process-as-a-service platform. Applied Sciences, 12(3), 1392.
15. Martínez-Cruz, N., Porcel, C., Tejeda-Lorente, A., & Herrera-Viedma, E. (2020). A recommender system to tackle enterprise collaboration: The Smart Canvas. Future Generation Computer Systems, 106, 246–263.
16. Restuccia, F., Das, S. K., & Payton, J. (2016). Incentive mechanisms for participatory sensing: Survey and research directions. ACM Transactions on Sensor Networks, 12(2), 13.
17. Konhäusner, P., Cabrera Frias, M. M., & Dabija, D.-C. (2021). Monetary incentivization of crowds by platforms. Információs Társadalom, 21(2), 101–119.
18. Omokaro, O., & Payton, J. (2014). Towards a framework to promote user engagement in participatory sensing applications (Technical report). University of North Carolina at Charlotte.
19. IBM. (2025). Data and AI governance: A complementary duo. IBM Think Insights.