TACJE

TRANSNATIONAL ACADEMIC JOURNAL OF ECONOMICS

Transnational Academic Journal of Economics, Year: 2025, Volume: 2, Issue: 3, Pages : 218-225 *Research Article*

DOI: 10.5281/zenodo.18442879

Publication Date: 30.09.2025

A. V. N. Murty[1]

[1] *KLEF Business School, KL University, India. Vaddeswaram, Guntur District, Andhra Pradesh 522502, India email dravnmurty@kluniversity.in ORCID: 0000-0002-5232-2226*

AI-Based Consumer Behaviour Modelling Under Algorithmic Transparency and Regulatory Constraints: A Governance-by-Design Framework Using the EU AI Act and Digital Services Act as Benchmarks

Abstract

AI-driven consumer behaviour modelling powers targeting, ranking, recommendations, dynamic pricing, and churn prediction, but it increasingly operates under legal requirements for transparency, risk management, and accountability. This paper develops a governance-by-design framework for non-EU jurisdictions by using the EU Artificial Intelligence Act (Regulation (EU) 2024/1689) and the EU Digital Services Act's recommender-system transparency orientation as comparative benchmarks. Drawing on the OECD Recommendation on AI and the NIST AI Risk Management Framework, we translate benchmark obligations into implementable lifecycle controls: data governance, model documentation, explainability, bias evaluation, audit logging, post-deployment monitoring, and incident response. To strengthen decision usefulness, we add a quantitative scenario layer that compares governance tiers over a 2026–2035 horizon on expected consumer-harm incidents, model performance retention under drift, and a regulatory-risk premium proxy. Results provide a modular control architecture, an implementation sequence, and metrics to reduce regulatory and reputational tail risks while preserving commercial effectiveness.

Keywords: consumer modelling; algorithmic transparency; EU AI Act; Digital Services Act; recommender systems; AI governance; explainable AI; risk management

How to cite: Murty A. V. N. AI-Based Consumer Behaviour Modelling Under Algorithmic Transparency and Regulatory Constraints: A Governance-by-Design Framework Using the EU AI Act and Digital Services Act as Benchmarks. *Transnational Academic Journal of Economics, 2*(3), 218-225

1. Introduction

Consumer behaviour modelling refers to the prediction and influence of consumer decisions using data-driven methods. In digital markets it is implemented through AI systems that infer preferences, propensities, and sensitivities to price, messaging, or product attributes. These models power recommender systems, personalization, advertising auctions, dynamic pricing, and customer lifecycle management. Because platforms increasingly coordinate market access and attention allocation, consumer modelling has become both a competitive capability and a regulatory focal point.Regulatory expectations are converging toward risk-based governance. The EU Artificial Intelligence Act establishes a horizontal framework for AI systems across risk tiers, with stricter duties for higher-risk systems (e.g., risk management, data governance, technical documentation, record-keeping/logging, transparency, human oversight, and robustness/cybersecurity). In parallel, the Digital Services Act requires online platforms to provide specific transparency about recommender systems and offer users at least one option that is not based on profiling. These benchmarks matter for non-EU jurisdictions because cross-border platform operations export compliance practices through product design and supply chains, while policymakers often reference EU rules when shaping emerging governance regimes.However, transparency is not self-executing. A disclosure that is not backed by evidence-producing controls can increase litigation and enforcement risk by creating a mismatch between claims and practice. Accordingly, this paper treats algorithmic transparency and regulatory constraints as design parameters and proposes a governance-by-design framework that is auditable, modular, and proportionate for emerging markets.Research objectives are to (i) identify risks relevant to consumer modelling (deception, discrimination, manipulation, opacity), (ii) translate benchmark requirements into implementable controls, (iii) propose a lifecycle governance architecture aligned to OECD and NIST principles, and (iv) add quantitative comparisons of governance tiers to support forward-looking planning.

2. Materials and Methods

Study design. We apply comparative governance analysis using three benchmark anchors: (i) the EU AI Act as the primary legal reference for risk-tiered obligations; (ii) the DSA recommender-system transparency orientation (including user choice requirements); and (iii) international governance standards (OECD AI Recommendation; NIST AI RMF) as implementable scaffolding. Mapping method. We map benchmark duties and trustworthy-AI principles into lifecycle controls across four stages: data, model, deployment, and oversight. Controls are evaluated by (a) risk coverage, (b) implementability, (c) auditability, (d) user meaningfulness, and (e) proportionality. Quantitative scenario layer. Because many jurisdictions lack public microdata on platform modelling outcomes, we add an illustrative scenario analysis that compares three governance tiers (baseline, governance-by-design, high-assurance) over 2026–2035. The scenario metrics are: (1) expected material consumer-harm incidents (count), (2) a regulatory risk premium proxy (percentage points) representing financing/insurance/legal friction attributable to governance weaknesses, (3) model performance retention under drift (index, 2026=1), and (4) governance/control operating cost as a share of AI operating expenditure. These scenarios are not an econometric estimate; they are structured planning comparisons intended to make the trade-offs explicit and testable.

3. Results

Figure 1 presents the governance-by-design lifecycle. Table 1 provides a control matrix mapping risk areas to implementable controls and audit evidence. To add decision-useful comparisons, Table 2 and Figures 2–3 provide an illustrative forward-looking scenario analysis of governance tiers (2026–2035).Figure 1. Governance-by-Design Lifecycle for Consumer Behaviour Modelling AI (Benchmark: EU AI Act + DSA)

DATA STAGE → MODEL STAGE → DEPLOYMENT STAGE → OVERSIGHT STAGE
Data: lawful collection, minimisation, provenance, representativeness checks, data lineage.
Model: model cards, evaluation/fairness/robustness, explainability plan by audience.
Deployment: user-facing disclosures and choice settings, logging and monitoring, human-oversight triggers.
Oversight: internal audit/assurance readiness, incident management, corrective actions, periodic risk reviews.

Table 1. Control matrix for consumer behaviour modelling AI: benchmark rationale, technical controls, and audit evidence

Risk / obligation area	Benchmark rationale (EU)	Implementable technical controls	Evidence artifacts for audit
User transparency (AI interaction)	Transparency obligations and guidance emphasis for AI-user interaction and information duties	AI interaction labels; 'why am I seeing this?' explanations; disclosure versioning; comprehension testing	UI logs; disclosure text versions; A/B comprehension results; help-centre records
Recommender-system transparency	DSA-oriented transparency for recommenders, including user choice mechanisms	Recommender explanation; main-parameter summaries; user settings (including non-profiling option)	Recommender documentation; user-setting telemetry; parameter change logs
Bias and discrimination	Trustworthy AI principles and risk frameworks emphasize fairness and non-discrimination	Bias diagnostics; subgroup evaluation; mitigation decisions; monitoring for disparate exposure/ outcomes	Evaluation reports; bias tests; mitigation approvals; monitoring dashboards
Manipulation / vulnerability exploitation	Benchmark regimes prohibit certain harmful practices and emphasize protection of vulnerable groups	Restricted targeting categories; safeguards for minors; vulnerability exploitation checks; policy enforcement	Policy rules; enforcement logs; incident reports; complaint analysis
Logging and monitoring	Record-keeping/logging and post-deployment monitoring expectations for accountable AI	Audit logging pipeline; drift and harm monitoring; alerting; incident response playbooks	Log-retention policy; dashboards; incident tickets; post-mortems
Governance alignment	OECD/NIST emphasize lifecycle governance, accountability, and continuous risk management	RACI matrix; model registry; periodic risk reviews; change control	Governance minutes; registry snapshots; release checklists; audit findings and closures

Table 2. Illustrative governance-tier scenarios (2026, 2030, 2035)

Governance tier	2026: incidents	2026: reg risk pp	2026: perf idx	2026: cost %	2030: incidents	2030: reg risk pp	2030: perf idx	2030: cost %	2035: incidents	2035: reg risk pp	2035: perf idx	2035: cost %
Baseline (minimal controls)	18.0	1.5	1.0	0.4	22.4	1.99	0.956	0.4	28.0	2.6	0.9	0.4
Governance-by-Design (moderate)	13.0	1.27	1.03	0.9	16.2	1.69	0.986	0.9	20.2	2.21	0.93	0.9
High-Assurance (risk-tiered + audit-ready)	9.9	1.12	1.05	1.4	12.3	1.49	1.006	1.4	15.4	1.95	0.95	1.4

Figure 2. Expected material consumer-harm incidents under governance tiers (illustrative projection)

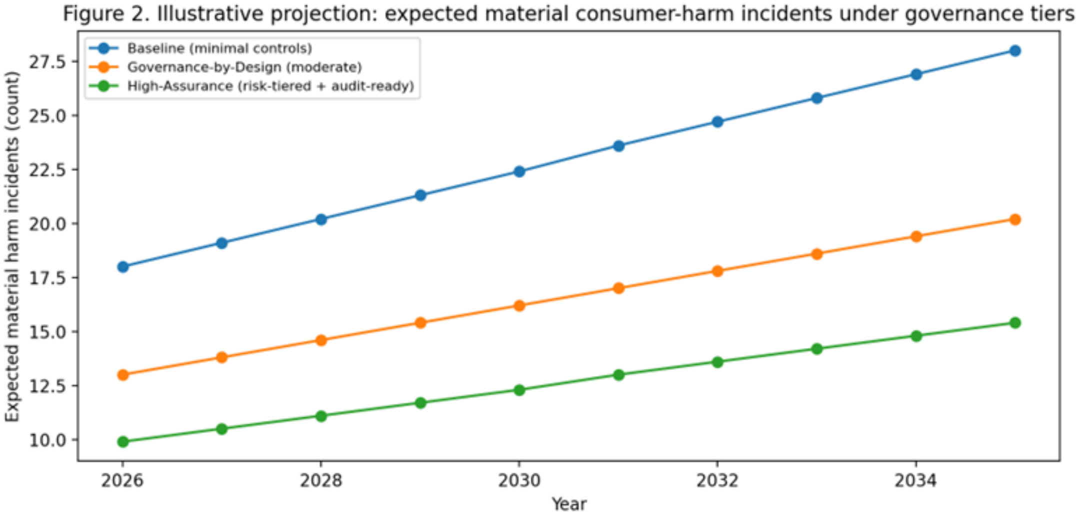
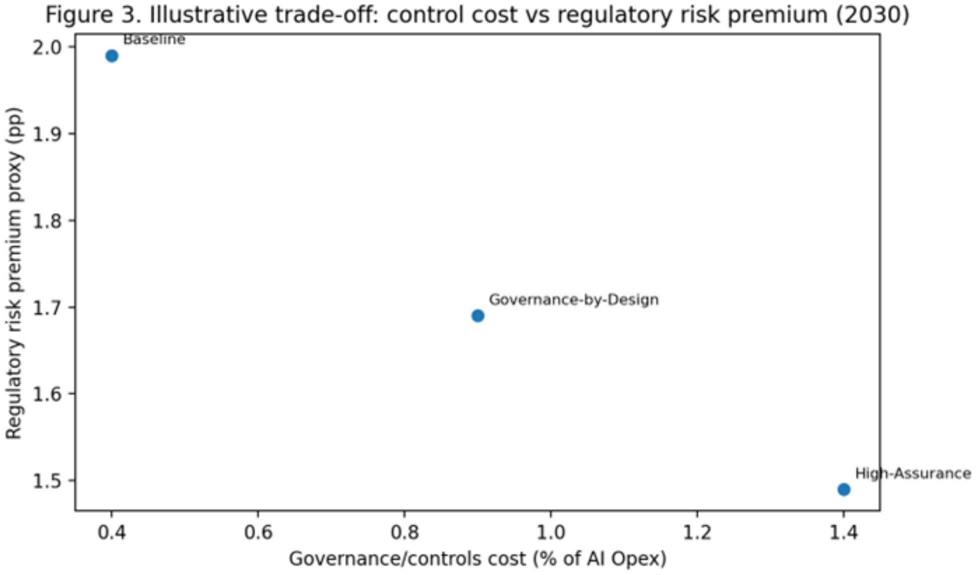


Figure 3. Trade-off between governance cost and regulatory risk premium proxy (2030, illustrative)



4. Discussion

The scenario layer highlights why governance-by-design is economically relevant. As platforms scale, the expected frequency of high-impact incidents rises under baseline controls because behavioural feedback loops, drift, and product experimentation increase exposure to opacity, discrimination, and manipulation risks. Governance-by-design reduces this risk primarily by (i) strengthening data lineage and evaluation discipline, (ii) creating evidence-producing transparency and user-choice mechanisms, and (iii) establishing monitoring and corrective-action loops. A key implication is that transparency should be treated as a control, not a communication artifact: disclosures must be traceable to model documentation, parameter settings, and monitoring evidence.For non-EU jurisdictions, proportionality is critical. The EU's risk-tiering logic can be translated into a two-speed approach: minimal baseline controls for low-impact personalization and stronger controls for high-impact systems such as dynamic pricing, eligibility-like gating, or targeting that affects vulnerable groups. This approach preserves innovation while materially reducing tail risks and regulatory friction.Finally, governance must manage an 'explanation leakage' risk: detailed explanations to end users can be exploited to game recommender or pricing systems. Role-based transparency is therefore recommended—investigator/auditor-grade evidence internally, and meaningful-but-safe user explanations externally.

5. Conclusions

AI-based consumer behaviour modelling is becoming a governed capability rather than a purely competitive asset. Benchmark regimes show that durable compliance and trust depend on lifecycle controls—data governance, documentation, logging, monitoring, and incident response—paired with user-facing transparency and meaningful choice. For emerging markets, a modular governance-by-design architecture can deliver disproportionate risk reduction if it prioritizes high-impact systems first and ensures that transparency claims are supported by auditable evidence.The quantitative scenarios provide a planning lens: modest increases in governance operating cost can be justified by reductions in expected harm events and a lower regulatory-risk premium proxy, especially over a 5–10 year horizon. Future empirical research should test these mechanisms using platform-level incident data, complaint records, and audit outcomes, and should evaluate how different transparency UI designs affect consumer understanding and welfare.

Patents

No patents are claimed. Potential patentable outputs would arise only from future proprietary implementations such as automated evidence generation, tamper-evident audit logging, or real-time fairness-drift monitoring systems.

Supplementary Materials

Supplementary materials may include a model registry template, model-card/datasheet templates, transparency UX testing protocol, fairness/drift monitoring specification, and an incident-response playbook for algorithmic harm events.

Author Contributions

Conceptualization, methodology, analysis, writing and editing, and visualization were performed by A. V. N. Murty.

Funding

No external funding was received.

Institutional Review Board Statement

Not applicable. The study uses public legal texts, standards and literature; no human participants were involved.

Informed Consent Statement

Not applicable.

Acknowledgments

The author acknowledges the availability of public EU legal texts and the contributions of OECD and NIST to trustworthy AI governance.

Conflicts of Interest

The author declares no conflicts of interest.

Appendix A

Minimum viable checklist: system inventory; data provenance/representativeness; documentation; evaluation (calibration + subgroup metrics); explainability plan; user transparency + choice; logging + monitoring; incident + corrective action governance.

Appendix B

Illustrative dashboard indicators: drift/performance decay; exposure parity; subgroup error rates; opt-out usage; incident rates; model update compliance; documentation completeness; vulnerability safeguards outcomes.

References

1. European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.

2. European Union. (2022). Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act). Official Journal of the European Union.

3. OECD. (2019). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). OECD Legal Instruments.

4. NIST. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology.

5. NIST. (2024). NIST AI RMF Generative AI Profile (NIST-AI-600-1). National Institute of Standards and Technology.

6. European Commission. (n.d.). European Centre for Algorithmic Transparency (ECAT). European Commission.

7. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you?: Explaining the predictions of any classifier. In Proceedings of KDD (pp. 1135–1144).

8. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems.

9. Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. In Advances in Neural Information Processing Systems.

10. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. MIT Press.

11. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54(6), 1–35.

12. Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In Proceedings of FAT* (pp. 59–68).

13. Sweeney, L. (2013). Discrimination in online ad delivery. Communications of the ACM, 56(5), 44–54.

14. Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. Proceedings on Privacy Enhancing Technologies, 2015(1), 92–112.

15. Shin, D. (2021). The effects of explainability and causability on trust in AI systems. International Journal of Human–Computer Studies, 148, 102551.

16. Kizilcec, R. F. (2016). How much information?: Effects of transparency on trust in an algorithmic interface. In Proceedings of CHI (pp. 2390–2395).

17. Zhang, Y., Lomas, D., & Koedinger, K. (2020). Explaining recommendations: Effects on user trust and performance. ACM Transactions on Interactive Intelligent Systems, 10(4), 1–31.

18. Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In Proceedings of FAT* (pp. 149–159).

19. Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. Polity.

20. Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press.

21. Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. American Economic Review, 105(5), 491–495.

22. Sunstein, C. R. (2016). The ethics of influence: Government in the age of behavioral science. Cambridge University Press.

23. Aridor, G., Che, Y.-K., Salz, T., & Zhao, Y. (2020). The economic consequences of data privacy regulation: Empirical evidence from GDPR. NBER Working Paper.

24. Cremer, J., de Montjoye, Y.-A., & Schweitzer, H. (2019). Competition policy for the digital era. European Commission.

25. Rexhepi, B. R., Rexhepii, F. G., Xhaferi, B., Xhaferi, S., & Berisha, B. I. (2024). Financial accounting management: A case of Ege Furniture in Kosovo. Quality – Access to Success, 25(200).

26. Daci, E., & Rexhepi, B. R. (2024). The role of management in microfinance institutions in Kosovo—Case study Dukagjini Region. Quality – Access to Success, 25(202).

27. Murtezaj, I. M., Rexhepi, B. R., Dauti, B., & Xhafa, H. (2024). Mitigating economic losses and prospects for the development of the energy sector in the Republic of Kosovo. Economics of Development, 23(3), 82–92.

28. Murtezaj, I. M., Rexhepi, B. R., Xhaferi, B. S., Xhafa, H., & Xhaferi, S. (2024). The study and application of moral principles and values in the fields of accounting and auditing. Pakistan Journal of Life and Social Sciences, 22(2), 3885–3902.

29. Rexhepi, B. R., Daci, E., Mustafa, L., & Berisha, B. I. (2024). Tax accounting in the Republic of Kosovo. Economics, Management and Sustainability, 9(3), 66–73. https://doi.org/10.14254/jems.2024.9-3.5