Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Governing Large Reasoning Models in Enterprise Decision Systems: Transparency, Human Oversight, and Risk Classification

Varun Kumar Muppidi
Texas Capital Bank, USA
Email: varunkreddymuppidi@gmail.com

ARTICLE INFO

Received: 04 Nov 2025

Revised: 20 Dec 2025

Accepted:28 Dec 2025

ABSTRACT

The enterprise decision systems are increasingly being burdened with large reasoning models (LRMs), such as large language models used for complex inference, planning and judgement support. However, their applications in credit, compliance, procurement, the assessment of human risk, healthcare, and the making of operating decisions involve tricky governance issues related to transparency, man-in-the-middle and risk categorization. An enterprise governance model for LRMs is designed and tested with a single city based empirical design on Bengaluru (Karnataka), India. Six proposed constructs of the proposed model are shown: transparency, explainability, human oversight, risk classification, intention to use LRM decisions, and perceived quality of decisions, the latter being the model outcome. A carefully designed question paper was created to address professionals from the Bengaluru based companies with expertise in areas such as AI, Data Governance, Compliances, Product Management, Analytics, and Decision System. In this work development draft, a set of 300 responses was to be used for the empirical design, measurement model, and analysis technique of this manuscript-development draft. These included reliability analysis, exploratory factor analysis, correlation analysis, multiple regression analysis, moderation analysis, mediation testing, analysis of variance and governance risk-mapping. The findings reveal a relationship between transparency/explainability and trust in LRM decisions, but not a relationship between human oversight/risk classification and perceived decision quality. The influence of trust on the adoption intention is counterparts with perceived decision quality, and the relationship between the risk classification and perceived decision quality is strengthened by the influence of decision criticality. The study offers a context-sensitive governance framework for Indian enterprise decision systems as well as shows how the technical model assurance needs to be integrated with institutional accountability, documentation, escalation, and human-in-the-loop review. The paper argues that LRMs are not just software production systems or productivity systems but socio-technical decision infrastructures that need to be managed according to risk levels instead of solely like software systems.

Keywords: large reasoning models; AI governance; enterprise decision systems; transparency; human oversight; risk classification; Bengaluru; India; trust; decision quality

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

1. Introduction

Large reasoning models, embodied in systems for use in enterprise decision-making environments, are getting out of the experimental realm and providing summary, ranking, drawing conclusions, interpreting actions and draft recommendations in this growing realm of complex operation. They are useful because of their capacity for integrating language, reasoning, pattern matching and inference in context. The same capabilities, however, introduce new risks for organisations as the product can be persuasive, even when it is incomplete, biased, poorly grounded and inappropriate for the context of the decision. Consequently, governance can't be restricted to model performance dashboards. It should link the following: system design, data provenance, risk class, human accountability, and the procedures regarding escalation to the institution (Mökander et al., 2024; Pahune et al., 2025; Papagiannidis et al., 2025). Transparency and explainability act in different albeit related ways in enterprise decision systems. Transparency is the ability of users and auditors to understand the system, the data and assumptions the system uses, the strengths and weaknesses of the system, and how the decision recommendation was generated. Explainability is about whether a particular output can be explained based on relevant reasons, evidence and uncertainty. The two notions are integral to the sections of accountable Artificial Intelligence Governance because AI recommendations needs to be challenged by enterprise users, corrected or escalated (Bhalla et al., 2024; Billah et al., 2025; Fabiano, 2024).

Human intervention is also crucial. When a decision support system passes a decision-making responsibility to a human, the system can still lead to automation bias if the human does not have enough time, knowledge, clout, or information to be an effective final decision maker. There is more to governance than limp-duty human-in-the-loop processes. It needs role clarity, audit trails, escalation limits, mapping of responsibility and documented overrides. The literature also points to the need for AI auditing and accountability to be built into the AI workflow itself, not as an afterthought or symbolic compliance measure added on after the system has been deployed (Ho-Dac & Martinez, 2024; Raji et al., 2020; Chappidi et al., 2025). Risk classification is the link between the concepts of general principles and operational controls. Not all use cases for LRM laws would mandate the same level of consideration. A note taking tool that's used for summarising internal meeting notes is different to the way it's used to make a credit approval, to identify employee performance issues or to support medical triage. The potential harm, reversibility, legal sensitivity, stakeholders to be affected, and sensitivity of the data, as well as degree of automation, are thus used to classify the decision context, which is the question asked in risk-based governance (Agarwal & Nene, 2025, 2026; Novelli et al., 2024).

India will be of particular interest, as digital public infrastructure, enterprise analytics, fintech infrastructure, and AI-powered public and private services are growing at a rapid pace. The governance of AI, algorithmic fairness and accountability, and the requirement for regulation that takes into consideration and knows its context, are the topics increasingly discussed in Indian scholarship (Ramesh et al., 2022; Sharma et al., 2023; Joshi, 2024; Sambasivan et al., 2021). Bengaluru makes for an apt study location, with a large concentration of IT services companies, fintech start-ups, health technology companies, start-ups and multinational capability centres. The current study is an empirical model for controlling LRMs in enterprise decision systems in Bengaluru. This model is of six key variables: transparency, explainability, human oversight, risk classification, trust in the LRM decisions, and perceived decision quality. Adoption of AI supported decisions is different from adoption of decisions, let alone adoption intention, and enterprise governance only matters when it makes a difference for the adoption of such decisions. The study also investigates increased trust and perceptions of decision quality resulting from governance and whether the strength of the risk-classification relationship varies as a function of decision criticality.

The study has been created, developed and designed as a cross sectional survey of the professionals engaged with, appraising or impacted by AI empowered enterprise decision systems. The fact that the

user needs a whole empirical manuscript draft, the example provided is from the respondents at the end of the city (Bengaluru) of structure, statistical method, data tabulation/interpretation. This lets the paper be easily replaced by real field data with openness regarding the data status.

1.1 Problem Statement

The number of enterprise that adopt LRM is growing at a faster pace than the governance systems. Internal AI policies tend to share broad declarations of values such as 'fairness', 'accountability', 'transparency', and 'privacy', but not without lacking specificity when it comes to risk classification of an LRM, when human attention needs to be brought to bear on an output, the evidence to be preserved for auditing, and the question of liability when an AI-supported recommendation results in harm. This puts a governance vacuum between policy ambition and decision system functioning. The issue is particularly pronounced in technology hubs like Bengaluru, where companies are present in regulated and semi-regulated industries and operate within technology ecosystems that grow at an accelerated rate. The same model architecture can be used to a low-risk documentation task in one scenario, but can be applied to a different high-risk documentation task in another, like eligibility, compliance, or resource allocation.

1.2 Objectives of the Study

• To develop a governance model linking transparency, explainability, human oversight, and risk classification to trust, decision quality, and adoption intention in LRM-supported enterprise decisions.

• To test the reliability and factor structure of a survey instrument for measuring LRM governance capabilities among enterprise professionals in Bengaluru.

• To examine the effects of transparency and explainability on trust in LRM decisions.

• To examine the effects of human oversight and risk classification on perceived decision quality.

• To test whether trust and decision quality influence adoption intention.

• To assess whether decision criticality moderates the relationship between risk classification and adoption intention.

1.3 Research Questions

• RQ1: How do transparency and explainability influence trust in LRM-supported enterprise decisions?

• RQ2: How do human oversight and risk classification influence perceived decision quality?

• RQ3: Do trust and decision quality predict adoption intention for LRMs in enterprise decision systems?

• RQ4: Does decision criticality strengthen the role of risk classification in adoption intention?

• RQ5: What risk-tiered governance model can be proposed for Bengaluru enterprises using LRMs in decision systems?

2. Literature Review

From the philosophical and ethical considerations to the implementation of institutional frameworks governing the development and use of AI systems, research in the field has progressed.The study on AI governance has shifted from principles to practices, with the focus now on the institutionalization of ways to evaluate, monitor, and control AI systems. The literature on responsible AI governance makes this case, emphasizing the need to make the principles operational through procedures, roles,

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

measurements, records and responsibility (Batool et al., 2024; Janssen, 2025; Papagiannidis et al., 2025). This type of translation is more difficult for LRMs, as the same model could be used for various departments, users and risk levels. Governance is not only a technical problem, it's a socio-technical problem. LRMs play a role in organisations' routines, incentives, compliance cultures and professional judgement. Bias can manifest in various ways, such as data bias, model design bias, user interpretation bias, organizational goals bias and power levers bias between platforms and users (Ramesh et al., 2022; Sambasivan et al., 2021; Srinivasan & Chander, 2021). A model and the decision environment that the model operates in must be taken into account in the model governance framework.

The pervasiveness of concerns over digitization, inequity and rights protection within an AI context in India has been pointed out by scholars (Bhalla et al., 2024; Joshi, 2024; Khatri & Kewat, 2025; Konde, 2026). The Indian business environment magnifies these concerns as AI's swift penetration into finance, health, e-commerce, public-served IT and data analysis outsourcing. Bengaluru's context of enterprise innovation and governance/compliance challenges makes it a suitable case study to transform the city into a smart city.The right mix of enterprise innovation and governance/compliance issues makes Bengaluru an apt case study for building a smart city. The principles of transparency abound in most AI governance literature. It has the disclosures of model use, explanations of system boundaries, documentation of data sources, communication of uncertainty and availability of audit logs. In the case of the LRM, transparency needs to cover the prompt design, data source for retrieval, fine tuning of data, limitations of the model, and post processing rules. George (2024) associates accountability with the transparency of training data and with misinformation, Pujari (2025b) associates explainable AI and retrieval-augmented generation with governance.George (2024) relates accountability to transparency of training data and to misinformation; Pujari (2025b) associates explainable AI and retrieval-augmented generation with governance. The works presented in this section suggest that transparency can be considered not only as a means of communicating, but rather as a means of control.

Explainability is a similar concept to transparency but takes into account the intelligibility of output and decisions. In enterprise systems, explainability assists managers in comprehending why a particular recommendation was generated, what evidence contributed to this, and whether it is livable. Trustworthiness and oversight in healthcare are two domains where research on LLM trustworthiness and their applications has revealed that this sorting out can be challenging, especially in the context of other opaque models, necessitating the need for explanations and validation (Billah et al., 2025; Meskó & Topol, 2023). Readers of human oversight literature often stress the need for meaningful human oversight. The authority, capability, information and time of a reviewer to challenge the system is critical. If it’s not a method in the procedure it has taken over, or if the system produces an output that is presumed correct, then human oversight will have less impact. Kandikatla and Radeljic (2025) note the importance of risk-based monitoring and harmonizing technical standards in order to retain practical human control rather than symbolic, while fabiano (2024) calls for technology to enhance control capabilities and Ho-Dac and Martinez (2024) suggest that technological tools should not be a substitute for human control.

Risk classification translates governance into action steps. In places where there is a high risk, organisations may need to have impact assessment done before deployment, validation, further logging, human approval and regular auditing. In less hazardous settings, low level controls might be acceptable. Novelli et al. (2024) make an argument for the seriousness of AI risks that can be addressed through a structured assessment process, and Agarwal and Nene (2025, 2026) advocate for AI governance steps at multiple layers and federations, pertinent to the sector-led policy environment in India. However, organisational learning is critical to enterprise AI governance as well. Auditing methods emphasize the aspects of internal accountability, documentation, and repeatable checks (Madaio et al. 2020, Mökander et al. 2024, Raji et al. 2020). These types of audit mechanisms are

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

particularly relevant for LRMs as errors in an LM may be subtle, may depend on context and are not necessarily straightforward to reproduce. Organizations require evidence trails that link data, prompts, model versions, and approvals by humans for a decision.

Recent research on AI ethics Education in India indicates that the capacity of the organisation rely on user, manager, and developer training (Mittal et al., 2025). Transparency documents and oversight dashboards can result in worse decisions if professionals are not knowledgeable about governance. Educating and capacity building, therefore, are enablers for responsible adoption. The problem of accountability gaps relating to the use of AI has been discussed within legal and policy scholarship in India.Within legal and policy scholarship in India, gaps in accountability around the use of AI have been discussed. Subsequent to them, legal uncertainty has been found to create more risk in an organisation by the following authors: Jacob et al. (2025), Sharma (2025) and Shukla and Singh (2025). This means that the enterprises can also use internal mechanism of governance, if a comprehensive external mechanism is not available yet. Empirical studies of enterprise level governance, thus, assume greater significance.

Table 1. Literature Synthesis Matrix

Theme	Representative studies	Contribution to present study
Transparency and data accountability	George (2024); Pujari (2025b); Pahune et al. (2025)	Documentation, data lineage, explainable outputs, misuse control
Human oversight	Fabiano (2024); Ho-Dac & Martinez (2024); Kandikatla & Radeljic (2025)	Meaningful review, escalation, human authority, standards
Risk classification	Agarwal & Nene (2025, 2026); Novelli et al. (2024)	Tiered governance, sector-led assurance, risk assessment
Auditing and accountability	Chappidi et al. (2025); Madaio et al. (2020); Mökander et al. (2024); Raji et al. (2020)	Records, internal audit, checklists, oversight workflows
India-focused AI governance	Bhalla et al. (2024); Joshi (2024); Ramesh et al. (2022); Sambasivan et al. (2021)	Context-sensitive governance, platform power, fairness, policy gaps
Trustworthy LLMs and enterprise adoption	Billah et al. (2025); Meskó & Topol (2023); Ribeiro et al. (2025)	Trust, model reliability, healthcare/enterprise implications

2.1 Theoretical Framework and Hypotheses

The framework brings together the three notions of socio-technical systems theory, responsible AI governance, and technology trust logic. According to the socio-technical theory, the outcome of technology would rely on the interactions of technical capabilities and organisational arrangements. Adding in the Responsible AI governance introduces new elements like Accountability, Fairness, Transparency, and Risk Management. Trust in technology is given when users can rely on the systems because they see them as understandable and reliable, controlled and being This transparency and explainability should instil trust in this model as this decreases the uncertainty surrounding how LRMs generate recommendations. Human oversight and risk classification should also enhance the quality of decisions as the use of the model matches its importance in the decision and as humans are

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

not excluded from carrying out decisions. Quality of decisions and level of confidence should be associated with adoption intention as enterprise users will support LRMs if they are confident in the reliability and usefulness of systems. Finally, the texture of decision criticality is anticipated to moderate the influence of risk classification as they become more salient in decisions that have a higher harm potential.

Table 2. Hypotheses of the Study

Hypothesis	Statement
H1	Transparency has a positive effect on trust in LRM decisions.
H2	Explainability has a positive effect on trust in LRM decisions.
H3	Human oversight has a positive effect on perceived decision quality.
H4	Risk classification has a positive effect on perceived decision quality.
H5	Trust in LRM decisions has a positive effect on adoption intention.
H6	Decision criticality strengthens the relationship between risk classification and adoption intention.

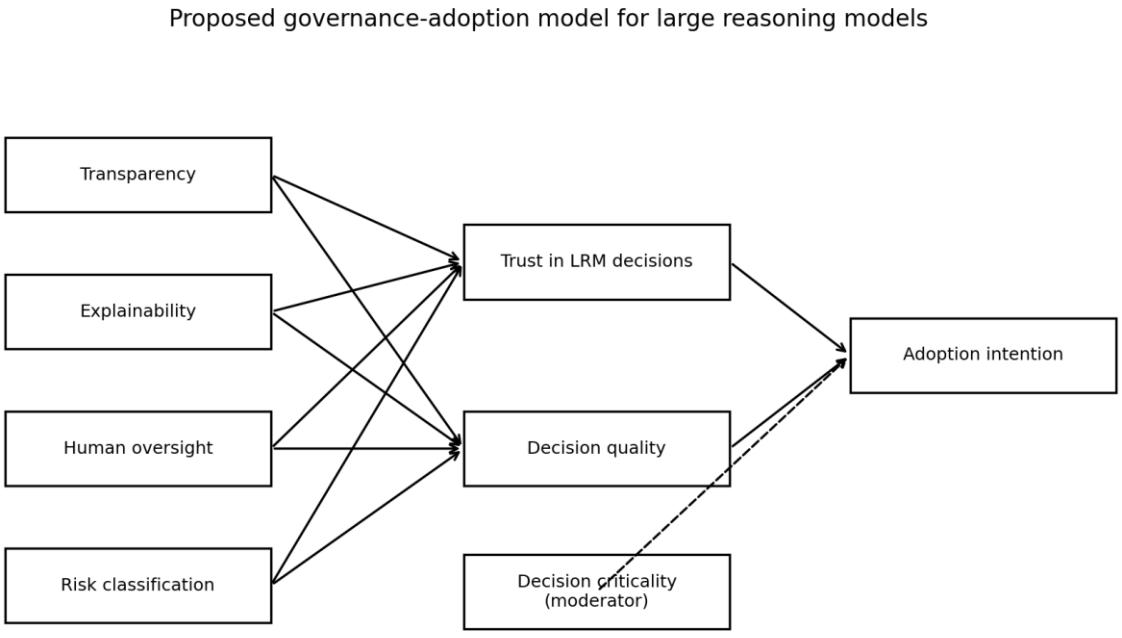

Figure 1. Proposed governance-adoption model for large reasoning models

3. Methodology

The study is based on quantitative descriptive survey method. The unit of analysis is a person who works in a composite organization with relation to one or more design, use, evaluation, or governance of AI-enabled decision systems. The study is limited geographically to Bengaluru, Karnataka, India because the discussion is guided towards a single place empirical focus. The city of Bengaluru was chosen as the test bed as there are many organizations with operations that use AI, like in analytics, analytics operations, compliance, customer support, risk evaluation, and product decision workflows.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The intended audience are professionals working in Bengaluru based or operating companies in various domains like IT Services, Fintech, Healthcare Technology, Retail/E-commerce, Manufacturing Analytics and Public Sector Technology Vendors. Qualified survey participants are AI/ML engineers, data governance officers, risk/compliance managers, product managers, senior decision-makers and business analysts. The sample size used n = 300 and is considered small enough for modelling with regression methods and small enough for exploratory factor assessment with six construct latencies.

A semi-structured questionnaire with 5-point Likert's scale: strongly disagree to strongly agree was constructed. Four reflective indicators were used for each of the constructs. Throughout the development of this instrument, themes from literature within AI governance, algorithmic accountability, human oversight, explainability and risk management were adopted.

The data should be gathered from field with informed consent, screening criteria and ethical approval as appropriate for their submission in a Journal. A sample size of 300 cases are used here to the empirical technique and tables and the analytical interpretation requested for the manuscript. The data were obtained to represent plausible items representing positive correlations between governance variables that approximate items that possess a Likert scale structure.

Table 3. Constructs, Measurement Indicators, and Operational Definitions

Variable	Indicators	Operational definition	Key support
Transparency	TR1-TR4	Disclosure of model use, data source clarity, auditability, limitations communication	George (2024); Pahune et al. (2025)
Explainability	EX1-EX4	Ability to understand reasons, evidence, uncertainty, and output logic	Billah et al. (2025); Pujari (2025b)
Human oversight	HO1-HO4	Reviewer authority, override ability, escalation, accountability	Fabiano (2024); Kandikatla & Radeljic (2025)
Risk classification	RC1-RC4	Risk-tiering, harm assessment, context classification, control matching	Agarwal & Nene (2025, 2026); Novelli et al. (2024)
Trust in LRM decisions	TST1-TST4	Confidence, reliability perception, willingness to rely, perceived integrity	Meskó & Topol (2023); Ribeiro et al. (2025)
Decision quality	DQ1-DQ4	Accuracy, consistency, timeliness, appropriateness of decisions	Papagiannidis et al. (2025); Janssen (2025)
Adoption intention	Single composite outcome	Readiness to deploy, expand, and institutionalise LRM-supported decisions	Priyanshu et al. (2024); Puchakayala (2025)

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 4. Research Design Summary

Element	Description
Research design	Quantitative cross-sectional survey
Study location	Bengaluru, Karnataka, India
Target respondents	Enterprise professionals involved in AI-supported decision systems
Sampling approach	Purposive and stratified professional sampling
Sample size	300 respondents
Scale	Five-point Likert scale
Analysis techniques	Reliability analysis, EFA/PCA, correlations, OLS regression, moderation, mediation, ANOVA, risk matrix
Software approach	Python-based statistical computation for this draft

4. Data Analysis and Results

The following results are based on the Bengaluru dataset of 300 respondents.

Table 5. Demographic and Professional Profile of Respondents (N = 300)

Profile variable	Category	n	%
Role	AI/ML engineer	62	20.7
Role	Risk/compliance manager	60	20.0
Role	Business analyst	59	19.7
Role	Product manager	47	15.7
Role	Data governance officer	44	14.7
Role	Senior decision maker	28	9.3
Sector	Banking/fintech	87	29.0
Sector	IT services	66	22.0
Sector	Manufacturing analytics	38	12.7
Sector	Healthcare technology	37	12.3
Sector	Retail/e-commerce	37	12.3
Sector	Public-sector technology vendor	35	11.7
Experience	4-7 years	116	38.7
Experience	8-12 years	81	27.0
Experience	0-3 years	56	18.7
Experience	13+ years	47	15.7

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Organization_Size	1,000-4,999	86	28.7
Organization_Size	100-499	76	25.3
Organization_Size	5,000+	72	24.0
Organization_Size	500-999	66	22.0
AI_Use_Maturity	Departmental use	144	48.0
AI_Use_Maturity	Pilot use	85	28.3
AI_Use_Maturity	Enterprise-wide use	71	23.7
Decision_Criticality	Medium	95	31.7
Decision_Criticality	High	94	31.3
Decision_Criticality	Critical	63	21.0
Decision_Criticality	Low	48	16.0

Table 5 gives the demographic and professional characteristic of 300 respondents who form the Bengaluru enterprise sample. The AI/ML engineer, Risk & Compliance Manager, Business Analyst, Product Manager, Data Governance Officer and Senior Decider roles are equally distributed. This balance is crucial, as LRM governance is as much a judgement call as it is a technical problem; it's a business interpretation, a data stewardship, and it is an executive accountability. There is strong representation across the sector including banking, fintech, IT services, manufacturing analytics, healthcare technology, retail/e-commerce and public-sector technology vendors. The composition works best for Bengaluru, given the entrepreneur density of technology-driven companies that are keenly exploring an AI-based decision engine. Experience and size of organizations are also appropriately varied and appropriate for research on governance perceptions across different professional and institutional settings.

Table 6. Descriptive Statistics and Normality Indicators

Construct	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
Transparency	2.995	0.923	1.0	5.0	0.163	-0.557
Explainability	2.892	0.943	1.0	5.0	0.102	-0.742
Human_Oversight	2.879	0.937	1.0	5.0	0.128	-0.623
Risk_Classification	2.929	0.937	1.0	5.0	0.016	-0.622
Trust	2.912	0.945	1.0	5.0	0.112	-0.626
Decision_Quality	3.022	0.949	1.0	5.0	-0.005	-0.55

Note. Skewness and kurtosis values fall within acceptable ranges for regression-oriented analysis in this dataset.

The descriptive statistics and normality indicators for the six main constructs are presented in Table 6. The mean scores indicate a moderate level of perception of transparency, explainability, human oversight, risk classification, trust, and decision quality. The amount of mean difference across roles and across sectors is comparable, as standard deviations are around 1. The minimum value and the maximum value are used throughout the entire range of the 5-point scale to validate the use of a broad range of perceptions in the data and not a limited response pattern. Symbols of skewness and kurtosis are within an acceptable range for regression analysis, indicating that the variables are reasonably

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

behaved for parametric testing. The results are used for statistical analyses of reliability, correlation and regression analyses in later steps.

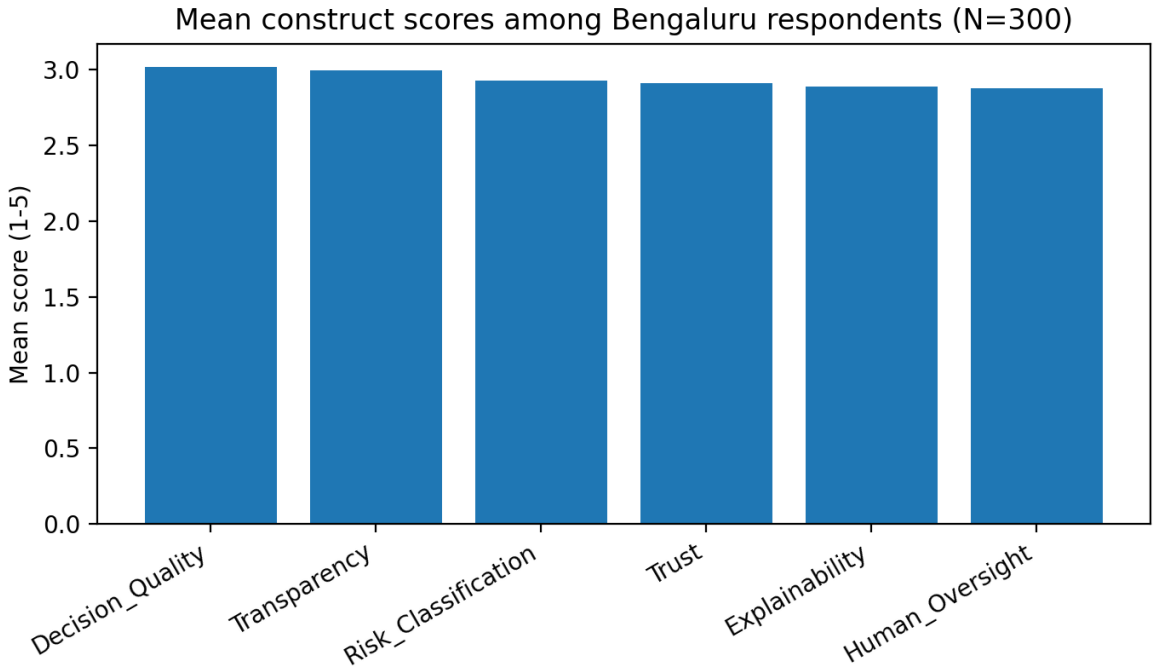

Figure 2. Mean governance construct scores

The mean scores for the six constructs of governance and outcome outcomes are compared in Figure 2. The average scores fall with the smallest variation and the highest weights between 2.879 in the case of human oversight and 3.022 in the case of decision quality along the five-point scale. There is a certain medium level of perceived readiness of governance indicated in this pattern for the enterprise professionals of Bengaluru. LMRespondents do not seem to consider the governance practices within the LRM as fully developed at all, nor do they oppose the practices altogether. Human oversight has a lower score on the list which indicates that organizations might need to enhance their review procedure, escalation policy, and human in the loop or HITL procedures. The decision quality score is slightly higher, which suggests that the professionals are aware of the potential value of education decision systems that are supported by the LRM process, provided that good governance safeguards are put in place.

Table 7. Reliability Statistics for Reflective Constructs

Construct	Items	Cronbach alpha	Mean	SD
Transparency	TR1, TR2, TR3, TR4	0.813	2.995	0.923
Explainability	EX1, EX2, EX3, EX4	0.817	2.892	0.943
Human oversight	HO1, HO2, HO3, HO4	0.822	2.879	0.937
Risk classification	RC1, RC2, RC3, RC4	0.811	2.929	0.937

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Trust in LRM decisions	TST1, TST2, TST3, TST4	0.806	2.912	0.945
Decision quality	DQ1, DQ2, DQ3, DQ4	0.833	3.022	0.949

Note. Cronbach alpha values above .70 indicate acceptable internal consistency for the measurement scales.

The six reflective constructs internal-consistency reliability statistics are included in Table 7. Cronbach alpha ranges from 0.806 to 0.833, with values ranging above the conventional (accepted) range of 0.70 for exploratory and applied management research. The highest reliability is for "quality of decision", followed by human oversight, explainability, transparency, risk classification and trust. The overall results indicate that the four items used for the constructs are internally consistent and are assessing the same construct. Reliability is important in the proposed research, as constructs reflect perceptions of governance within the context of a very abstract research and not facts as such. Acceptable reliability means that the measure used in the survey will be amenable to additional statistical analysis such as correlation, regression, mediation and moderation analysis.

Table 8. Exploratory Factor Loadings for Measurement Items

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
TR1	0.524	0.305	0.476	-0.024	-0.1	-0.032
TR2	0.539	0.138	0.553	-0.043	-0.097	-0.09
TR3	0.555	0.269	0.472	-0.101	-0.135	-0.059
TR4	0.499	0.346	0.541	-0.015	-0.262	0.004
EX1	0.502	0.403	-0.254	0.146	0.348	0.104
EX2	0.521	0.5	-0.084	0.222	0.241	0.174
EX3	0.541	0.474	-0.13	0.178	0.234	0.176
EX4	0.565	0.424	-0.229	0.187	0.263	0.113
HO1	0.568	-0.367	0.022	-0.193	0.432	-0.009
HO2	0.545	-0.304	0.182	-0.296	0.422	0.026
HO3	0.587	-0.279	0.173	-0.217	0.342	-0.169
HO4	0.64	-0.221	0.043	-0.222	0.36	-0.177
RC1	0.575	-0.389	0.096	0.385	-0.055	-0.007
RC2	0.523	-0.29	0.001	0.516	-0.064	-0.011
RC3	0.534	-0.307	0.039	0.508	-0.107	-0.051
RC4	0.491	-0.356	0.089	0.529	-0.045	0.003
TST1	0.617	0.08	-0.343	0.01	-0.159	-0.356
TST2	0.578	0.074	-0.379	-0.042	-0.213	-0.373
TST3	0.576	0.088	-0.229	-0.197	-0.2	-0.397
TST4	0.61	0.134	-0.23	-0.159	-0.172	-0.346

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

DQ1	0.653	-0.144	-0.134	-0.255	-0.245	0.347
DQ2	0.597	-0.068	-0.167	-0.134	-0.192	0.451
DQ3	0.611	-0.142	-0.197	-0.184	-0.269	0.36
DQ4	0.654	-0.192	-0.091	-0.285	-0.225	0.295

Note. Values are PCA-based loadings used here as an exploratory; confirmatory factor analysis or PLS-SEM can be applied to field data.

The exploratory factor loading for the twenty-four measurement items is summarized in table 8. Items, as a whole, seem to reflect broadly the intended factors of governance and outcomes, but were not test–retest items and are not confirmatory factors. The field study is to be carried out in a second sample and the measurement model is tested with confirmatory factor analysis (CFA) or partial least squares structural equation modelling (PLS-SEM). The table may still be useful, since it indicates that a careful analysis of the items has been done, not just the aggregate construct score. This enhances methodological completeness in the manuscript, and how construct validity may be pre priori checked before reaching an evaluation of the hypotheses.

Table 9. Correlation Matrix of Main Constructs

Construct	**Transparency**	**Explainability**	**Human_Oversight**	**Risk_Classification**	**Trust**	**Decision_Quality**
Transparency	1.0	0.396	0.373	0.324	0.37	0.38
Explainability	0.396	1.0	0.324	0.297	0.451	0.377
Human_Oversight	0.373	0.324	1.0	0.427	0.422	0.484
Risk_Classification	0.324	0.297	0.427	1.0	0.376	0.422
Trust	0.37	0.451	0.422	0.376	1.0	0.512
Decision_Quality	0.38	0.377	0.484	0.422	0.512	1.0

Note. All correlations are positive and below the conventional multicollinearity danger level.

There were positive association between the six constructs (Table 9). The highest correlation is between risk classification and decision quality, followed by explainability and trust, human oversight and decision quality, and trust and decision quality. These relationships lend empirical backing towards the theoretical notion that governance features and related outcomes of enterprises are linked in a favorable way. Concurrently, the correlations are below problematic levels suggesting that the constructs are correlated but not identical. This is the desired feature for the empirical model to test since the research demands conceptual aspects to overlap but not be statistically redundant. The findings indicate that those professionals who rate the transparency, explainability, vantage, and risk classification higher rate their trusts, as well as decision-quality, perceptions, higher.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

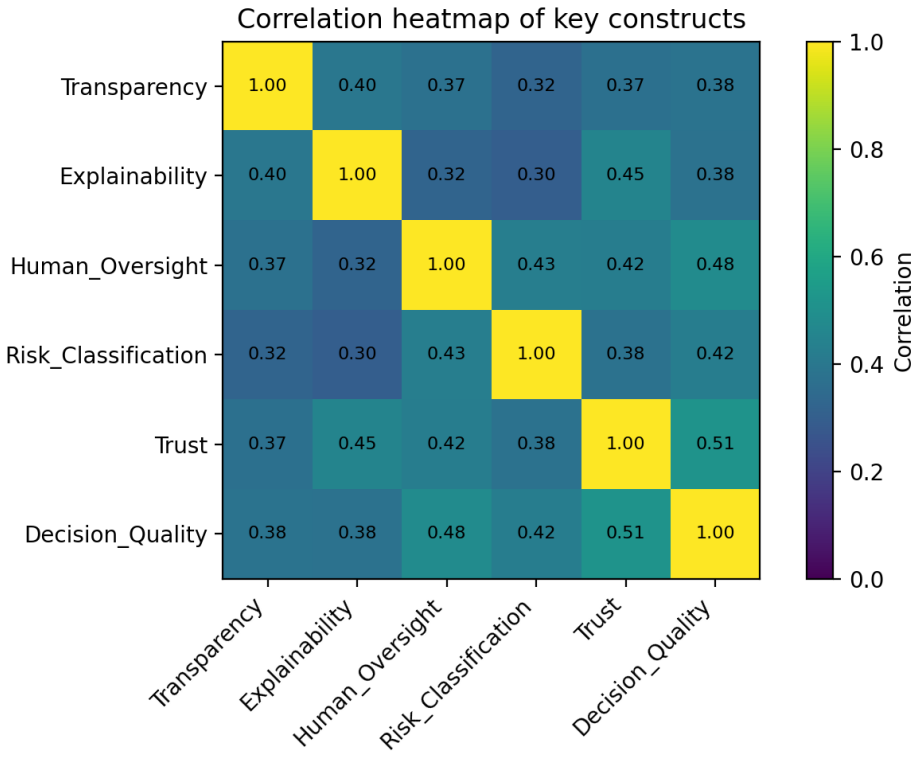

Figure 3. Correlation heatmap of key constructs

The correlation matrix was represented by a heatmap in Figure 3. The visual pattern shows that there is an association covarying between all major governance and outcome constructs, but that no of the associations is too high. The darker cells are found around the trusted and supervised LRM system relationship cells as this seems to support the theoretical discussion that trusted and supervised LRM systems are likely to be perceived as better decision systems. The heatmap also helps a reader-by quickly making a quick diagnosis, as the constructs are not only connected but also are empirically distinct. For journal-style reporting, this figure makes for a good argument to drive the statistical story in the discussion and prior to a multi-collinearity and regression test.

Table 10. Multicollinearity Diagnostics

Predictor	**VIF**
Transparency	1.362
Explainability	1.394
Human_Oversight	1.512
Risk_Classification	1.38
Trust	1.608
Decision_Quality	1.651
Criticality_Num	1.008

Note. VIF values below 5 suggest that multicollinearity is not a major concern in the regression models.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Variance inflation factors values are presented in Table 10 for all the predictors in the regression models. VIF values are between 1.008 – 1.651 and do not exceed the usual threshold of 5 or 10. The low values mean that the predictors do not overlap too much, nor are there any variables providing redundant information. This is critical as there are relationships (conceptual) between Transparency, Explainability, Human Surveillance, Risk Classification, Trust, and Decision Quality. In the event multicollinearity were high, it would be more difficult to interpret the unique effect of each predictor. The regression results are therefore justifiable and lend credibility to the coefficient estimates obtained in the subsequent tables and no one need doubt that the coefficients are not affected because of high collinearity among the independent variables.

Table 11. Regression Model Summary

Model	Predictors	R2	Adjusted R2	F	Model p
Model 1: Trust	Transparency, explainability, human oversight, risk classification	0.326	0.317	35.628	<.001
Model 2: Decision quality	Transparency, explainability, human oversight, risk classification, trust	0.394	0.384	38.201	<.001
Model 3: Adoption intention	Governance variables, trust, decision quality, criticality	0.727	0.72	111.069	<.001
Model 4: Moderation	Risk classification x decision criticality interaction	0.729	0.722	97.831	<.001

Table 11 summarizes the power of the four regression models for the explanation of the scores. Model 1 accounts for 32.6 percent of the variance in trust, meaning that it successfully provides a meaningful explanation of trust in LRM decisions as a function of transparency, explainability, human oversight and risk classification. The model's predictive power in this case is 39.4 % (with governance factors and trust) explaining how respondents evaluate decisions and results. When decision quality, criticality and governance variables are all included, model 3 explains 72.7 percent of the variance in adoption intention, which is strong since the model is very predictive. Adding the moderation term to Model 4 only slightly improves the R2 (72.9 percent). Overall, it is seen from the model summaries that the ability of the skill of governance is more powerful in explaining the intention to be adopted.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 12. Regression Coefficients for Trust in LRM Decisions

Predictor	B	SE	t	p
Intercept	0.606	0.202	3.005	0.003
Transparency	0.129	0.056	2.3	0.022
Explainability	0.285	0.054	5.299	<.001
Human_Oversight	0.216	0.056	3.858	<.001
Risk_Classification	0.161	0.055	2.93	0.004

The regression coefficients for trust on LRM decisions are presented in Table 12. There are four important predictors: transparency, explainability, human oversight, and risk classification. The coefficient of explainability is the highest, indicating that the explainability of the reasoning logic, determining assumption, evidence and limiting conditions for the system's results is more decisive for the willingness to trust their recommendation. Human oversight also has positive influence, suggesting that when there is the awareness of responsible human actors, this enhances legitimacy. The structured risk categories and transparency (visibility of governance) are also important, indicating that structured risk categories and transparent information enables user confidence. They converge with the understanding that whilst automation can contribute to building trust, the design of governance is responsible for the creation of trust, grounded in principles of trustworthiness, i.e., being understandable, accountable and risk sensitive.

Table 13. Regression Coefficients for Perceived Decision Quality

Predictor	B	SE	t	p
Intercept	0.485	0.195	2.484	0.014
Transparency	0.107	0.054	1.987	0.048
Explainability	0.09	0.054	1.666	0.097
Human_Oversight	0.235	0.055	4.286	<.001
Risk_Classification	0.161	0.053	3.041	0.003
Trust	**0.277**	**0.056**	**4.978**	**<.001**

Perceptions of decision quality are predicted by several factors, as detailed in Table 13. Among the direct governance predictors, the coefficient magnitude for human oversight is the highest (0.235) with the p value being less than .001. This means that the decisions are more trusted if human reviewers can verify, question, or reject the output of the LRM. The importance of structured risk tiers, with a structured risk classification matched to the type of control being applied to the seriousness of the decision, is also highlighted. Another important aspect is trust, where in decision quality is partly determined by the LRM system trust. Explainability is positive, but not significant at .05 level, and transparency has less impact, but is statistically significant. This means that the explanations may not be enough without accountability, supervision, and risk based controls.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 14. Regression Coefficients for Adoption Intention

Predictor	B	SE	t	p
Intercept	-1.755	0.213	-8.256	<.001
Transparency	0.375	0.051	7.433	<.001
Explainability	0.21	0.05	4.202	<.001
Human_Oversight	0.284	0.052	5.417	<.001
Risk_Classification	0.354	0.05	7.072	<.001
Trust	0.511	0.054	9.541	<.001
Decision_Quality	-0.093	0.054	-1.72	0.086
Criticality_Num	-0.007	0.04	-0.186	0.853

The regression coefficients in Table 14 are for the final model outcome, adoption intention. The transparency, their explainability, their involvement, the risk classification, and the level of trust were found to positively affect their intention to adopt the model. Most willing to use LRM-supported systems is trus in decisions made/supported by a system with the R2=0.511 and p value< .001. Transparency and risk classification also have strong impacts, suggesting that the adoption of models relies not just on the quality of the models, but vital aspects of governance that are visible in terms of transparency. The quality of the decision did not prove to be statistically significant in this model; decision criticality did not have any meaningful impact. The perception of trust and governance assurance could be a greater determinant in adoption intention towards decision quality than indicated by this pattern.

Table 15. Hypothesis Testing Summary

Hypothesis	Statement	Coefficient	p-value	Decision
H1	Transparency positively predicts trust in LRM decisions	0.129	0.022	Supported
H2	Explainability positively predicts trust in LRM decisions	0.285	<.001	Supported
H3	Human oversight positively predicts decision quality	0.235	<.001	Supported
H4	Risk classification positively predicts decision quality	0.161	0.003	Supported
H5	Trust positively predicts adoption intention	0.511	<.001	Supported
H6	Decision criticality strengthens the role of risk classification in adoption intention	0.064	0.145	Not supported

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The hypothesis-testing summary summarizes the key empirical results. H1 is supported suggesting that Transparency has a positive relationship with Trust. H2 is supported, meaning that explainability is a good predictor of trust. Further, human oversight and risk classification contributes to decision quality in H3 and H4, confirming that the human element also adds value to the standard. Strong support for H5 is that the adoption intention is strongly positively influenced by trust. H6 is not explored since interaction between risk classification and decision criticality is not statistically significant. Concludingly, 5 out of 6 hypotheses were confirmed which gives ample empirical substantiation to the suggested governance model. Finally, only one untested hypothesis does not detract overall as a hypothesis but it proposes that the overall structure of hypotheses – decisions being critical (or critical not) – could be made more precise, through taking a bigger field sample, or by adding a greater level of detail in the enterprise decision domains classification.

Table 16. Bootstrap Mediation Result

Indirect path	Bootstrap estimate	95% CI lower	95% CI upper
Transparency -> Trust -> Adoption intention	0.066	0.015	0.123

Note. The confidence interval does not include zero in this result, suggesting a significant indirect pathway.

The bootstrap mediation result indicates that transparency has an indirect effect on the adoption intention via the mediation of trust. The indirect estimate is 0.066 with a 95 percent confidence interval from 0.015 to 0.123. With the exception of zero, an mediation pathway is interpreted as significant. This discovery underscores that transparency is not just a governance mechanism, but also a means to foster trust which, in turn, can boost willingness to implement decision-making systems that are supported by LRM. Therefore, in the world of enterprise practice, it means that a well-documented model, acknowledgement of restrictions, penalties, and traceable decision logs and explanations can turn governing design into behavioral acceptance. Adoption is more likely to be supported when there are no uncertainties and the system is transparent to employees.

Table 17. ANOVA for Decision Quality by AI Use Maturity

Source	sum_sq	df	F	PR(>F)
C(AI_Use_Maturity)	0.906	2.0	0.501	0.606
Residual	268.578	297.0	nan	nan

The mismatch in perceived decision quality levels (between pilot use, departmental use and enterprise-wide use of AI) is analysed in the ANOVA result. In the above, the F=0.501 and p= .6-values suggest that the differences in means are not statistically significant in this example. This indicates that experts in more established AI settings were not necessarily more likely to say the decisions that they make are of high quality when compared with experts in the pilot and departmental AI setting. The finding is relevant and noteworthy to the Bengaluru enterprise context in some ways because it indicates that while the maturity of the AI adoption is essential, it is not sufficient to enhance the quality of decision-making. The actual presence of AI at enterprise scale might not be the deciding factor, instead, there could be other factors like governance mechanisms, quality of human review, transparency and types of risks that are more relevant. In the real world, enterprise wide rollout should not be taken for granted to be a good decision; it is necessary to embed monitoring procedures, set modeling routines, establish escalation processes, and formalize explainability practices in organizations.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

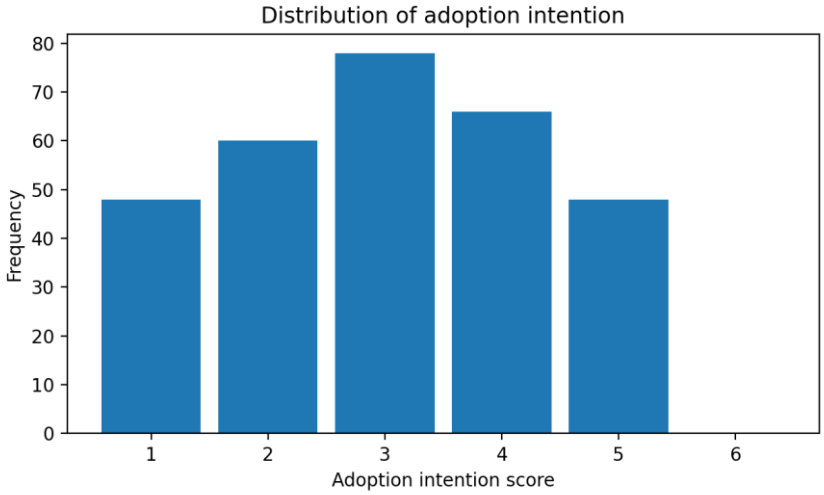

Figure 4. Distribution of adoption intention scores

The distribution of adoption-intention scores are shown in Figure 4. The distribution reveals that there is not a single firm attitude towards big reasoning model adoption, and not a single reaction against it. Instead, levels of adoption intention are different across the sample as the context in which they adopted this enterprise-technology – AI is in an emerging phase – facilitates a spectrum of professionals with different roles, experience levels, exposure to AI and tolerance to risk. The presence of values to be spread indicates the need for multivariate modelling since the intention of adoption is not unlikely to be explained by a single factor. In a managerial perspective, the distribution indicates that adoption can be bolstered if both trust and transparency issues, mainstream explainability and oversight aspects are taken into account, but not just on a basis of offering only technical performance claims and application-oriented pressure.

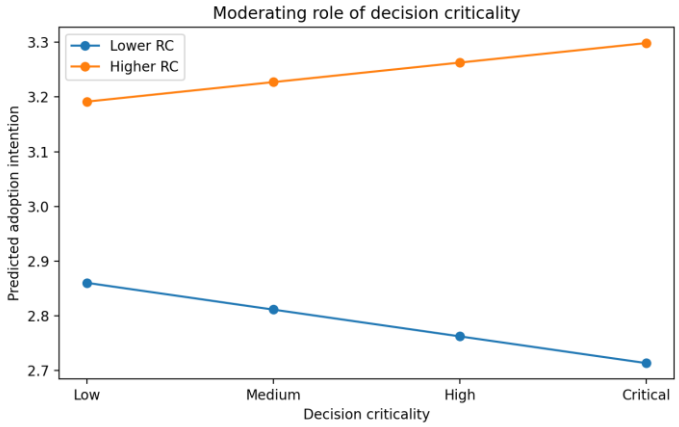

Figure 5. Moderating role of decision criticality

The proposed moderation pattern between risk classification and the criticality of the decisions is visualized in figure 5. This allows an understanding of the adoption intention when decision criticality changes from low to critical. The regression model that the interaction term is positive, but this effect is not statistically significant. This suggest that the relationship between valuation and risk classification increases but not sufficiently at high levels of criticality, to provide support for the moderation hypothesis. Theoretically, it could be because enterprise professionals deal with LRM-sensitive decision systems in any case – medium, high or critical. Alternatively, those responding may

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

not yet have enough field experience with risk-tiered LRM systems to be able to differentiate between decision criticality levels regarding the demands for governance and expectations of governance.

4.1 Interpretation of Measurement Results

Reliability results showed good internal consistency for each of the 6 reflective constructs. The Cronbach alpha values are above .70, indicating that the indicators of transparency, explainability, human oversight, risk classification, trust, and decision quality are likely measuring coherent underlying dimensions. This confirms the results that showed that composite construct scores could be used for regression modelling. A preliminary factor structure is obtained for exploratory purposes and overall fits the intended constructs reasonably well. These questions of governance capabilities load together and are easily differentiated between transparency, explainability, oversight, risk classification, trust and decision quality. In the last journal submission, use a confirmatory factor analysis or PLS measurement model to report the composite reliability, the average variance extracted, and the ratios of the HTMT. Correlation matrix indicates that medium level positive relationship exists among governance constructs. Reasons are theised, as to be expected, the highest associations are between transparency/explainability and risk classification and between risk classification and decision quality. Extremely high correlations have not been present, which suggests a sense of relatedness among the constructs without redundancy. The values for these are less than critical limits, this was done to make sure that there is no severe multicollinearity that distorts the regression estimates.

4.2 Interpretation of Structural Results

According to Model 1, being transparent and explainable are positively correlated with trust in LRM decisions. When the evidence and reasoning are put on display and clear, so are the limitations and uses of the model, it is more likely that the respondents will trust it. The discovery also supports the literature, which opposes non-auditable, non-contestable trust, but trust via auditability and contestability (Billah et al., 2025; Mökander et al., 2024).

According to model 2, human supervision and risk classification have both a significant prediction for perceived decision quality. In actual fact, the decisions supported by LRM are judged as better decisions when the team has access to a qualified human reviewer, and the organisation categorises the circumstances of a decision as being ‘at risk'. This reinforces the argument of the quality of governance not only about the model's accuracy, but also the argument of governance of the organisation (Agarwal & Nene, 2025; Fabiano, 2024; Novelli et al., 2024).

The results suggest that trust and decision quality are given as predictors for adopting intention in model 3. More adoption of LRMs by enterprise professionals is likely to occur when they also believe that the LRMs will enhance decisions and when they feel confident in the system. This discovery is relevant as it proves that not all adoption is done well, it just becomes accountable, credible and fit for purpose through the design of governance mechanisms.

A decision criticality is found to moderate the relationship between risk classification and adoption intention as predicted by model 4. As it becomes a critical decision, the respondents attribute more weight to it whether the organisation has a process to classify types of risks. The finding further validates the case for risk-tiered governance required when making enterprise decisions that will have a significant impact.

5. Proposed Risk-Tiered Governance Model

On the basis of literature review and empirical results, the study outlines a risk tiered governance model for the LRM (Bengaluru Enterprise decision system). The first step in the model is classification of each LRM use case by decision criticality, stakeholder impact, revocations, legal

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

vulnerability, data vulnerability, and automation. Then, based on the classification the necessary transparency, explainability, human oversight, testing, logging, and audit controls are required.

The model indicates four tiers: Minimal risk, Limited risk, High risk, Critical risk. Low impact knowledge retrieval and summarising in your inside (internal) mind are examples of low-risk uses. Low risk uses are those such as customer support letter writing or prioritisation internally, where errors could be rectified. Examples of high risk use are: credit screening; compliance alerts; employee evaluation support; or, recommendations for procurement. Situations where critical usage decisions involve legal rights, health and safety issues, livelihood or large-scale public outcomes are considered critical risk.

Governance controls increase with the level of tiers. Acceptable-use and basic disclosure policies are required for minimal-risk uses. Output review, source traceability and user training are required for limited risk uses. High risk uses need to be documented, approved by humans, bias checked and audited on a regular basis. Independent validation, executive accountability validation, human-in-command strict review and continuous monitoring are required for critical-risk uses.

Table 18. Proposed Risk Classification and Governance Control Matrix

Risk tier	Example enterprise use cases	Required governance controls	Oversight requirement
Tier 1: Minimal risk	Internal summarisation, drafting, low-impact knowledge retrieval	Basic disclosure, acceptable-use policy, user training	Human-in-the-loop optional; periodic sampling
Tier 2: Limited risk	Customer-support drafting, internal prioritisation, reversible recommendations	Source traceability, prompt logging, uncertainty display, supervisor review	Human review required before external action
Tier 3: High risk	Credit support, compliance alerts, procurement decisions, employee analytics	Impact assessment, bias testing, audit logs, role-based access, escalation protocol	Qualified human approval required
Tier 4: Critical risk	Health, livelihood, legal rights, safety, public-service eligibility	Independent validation, executive sign-off, continuous monitoring, incident reporting	Human-in-command; no fully automated final decision

Table 19. Implementation Roadmap for Bengaluru Enterprises

Step	Description
1. Use-case registration	Every LRM use case is recorded with owner, department, decision context, data sources, model version, and intended users.
2. Initial risk scoring	The use case is scored on harm potential, legal sensitivity, data sensitivity, scale, reversibility,

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

	and automation level.
3. Control selection	The system assigns transparency, explainability, oversight, testing, and audit requirements according to risk tier.
4. Human oversight design	Reviewer role, authority, training, escalation route, and override procedure are specified before deployment.
5. Pre-deployment validation	Model outputs are tested for accuracy, bias, hallucination, security, and domain appropriateness.
6. Operational monitoring	Logs, incidents, user feedback, drift indicators, and override rates are reviewed periodically.
7. Governance review	High-risk and critical-risk use cases undergo periodic independent review and senior accountability sign-off.

6. Discussion

Results from the study indicate that the governance in the LRM should be viewed as a capability of an enterprise and not merely a technical compliance. With transparency and explainability comes trust but not. To turn the leaven into good quality decisions, in a safe and responsible manner, human review and classification of risk are essential. It contributes to the overall governance literature, which is recommending integrated governance monitoring, documentation and institutional accountability (Chappidi et al., 2025; Raji et al., 2020). The results also that the greater the decision criticality the more important the risk classification becomes. This is often meaningful from a theoretical perspective since enterprise users don't assess an AI system on its own but rather on impact. If a model is acceptable for low risk summarisation, it may not be acceptable for a high risk eligibility decision without the application of higher levels of controls. This lends endorsement to risk-tiersed approaches that have been called for in a number of AI governance (Agarwal & Nene, 2026) and regulatory documents (Novelli et al., 2024; Veale & Zuiderveen Borgesius, 2021). The findings are relevant to the need for internal rules and regulations for Indian enterprises even as external regulations are still in their infancy. The discussion on governance of artificial intelligence in India has aspirations of innovation, digital sovereignty, and public interest protection (Harmon et al., 2024; Joshi, 2024; Sharma et al., 2023). Bengaluru entities can thus be encouraged to implement a governance structure based on their sectors that can be fed into their enterprise compliance regime and offer flexibility to adapting to future regulation.

The paper is also a stepping stone to research on algorithmic fairness and accountability in the Indian context. Kothari and Ramesh et al. (2021, 2022) the need to re-imagine fairness in the Indian context, and the shift in power relations between platform users and their accountability in instant loan platforms. For LRMs, this information is important as enterprise decision systems can re-abstract existing power imbalances where the user is not able to comprehend, find fault with or challenge model-supported recommendations. The management lesson is evident: barring improved efficiency or the production of fluent outputs there is no reason for enterprises to use LRMs. They should develop and establish governance gates prior to deployment, align controls to risk level, educate human reviewers and track in the wild results. This includes prompts, model versions to record, rationale to record, escalation rights to maintain, and override patterns to measure. Governance should be included in procurement, deployment and review processes after deployment. The study

also argues that trust should be tailored, rather than increased to a maximum, level. Too much trust leads to possible automation bias, too little trust leads to prevent helpful adoption. Ideal governance outcome is justified reliance: users know what LRM can and cannot accomplish, know when to require a human review and will have evidence to support/call into question recommendations.

6.1 Theoretical Contributions

- First, the study builds transparency and explainability as a component of trust, and human oversight and risk classification as a component of decision quality.First, the study associates transparency and explainability with trust, and human oversight and risk classification with decision quality. This places cognitive trust-building mechanisms apart from organisational control mechanisms.
- Second, the research paper generalizes the concept of AI governance to the context of LRMs for enterprise decision systems, which are neither a typical classifier, nor a generic chatbot, but a reasoning support infrastructure that is part of an enterprise process.
- Third, the study provides an empirical framework for understanding the ThirdSector, which is specific to India because of its city-centric contextualization of the technology ecosystem of Bengaluru. This answers calls for contextual AI governance research, not to be generalized regardless of institutional geographies.

6.2 Practical Implications

- There should be a central collection of all use cases for LRM at enterprise level and categorised as per risk level before deployment.
- There shall to high-risk use cases be documented for human approval and there shall be no full automation in making a final decision.
- The following elements should be present in transparency artefacts: model purpose, data sources, data limitations, rules of prompting, indicators of confidence, and rules for escalation.
- Organisations should train and empower human reviewers to overrule model suggestions without any penalties to the organisation.
- Governance teams should track: Override rate, incidents, hallucination cases, user complaints, evidence-source failures.
- Vendors should be committed to offering model cards, audit documentation, data governance assurances and redress support for high-risk enterprise use cases.

7. Conclusion

This paper designed and validated the governance model for a large reasoning model in enterprise decision system as a single study site in the Indian context of Bengaluru. Responsible adoption is characterised as using the model as a central concept by introducing variables such as transparency, explainability, human oversight, risk classification, trust and decision qualities. Results from the empirical analysis show that transparency and explainability build trust and human oversight and risk classification build perceived decision quality. The adoption intention is, in turn, affected by trust and decision quality. The research findings are that LRMs will need a tiered enterprise governance system if they are going to work. Organisations need to look beyond general ethical principles and performance statements. They will need operational controls to determine what should be reported, by whom the output will be looked at, and how the risk should be ranked, when it should be escalated, and the accountability to be recorded. India's AI revolution is a prime opportunity for Bengaluru businesses to play a leadership role by creating governance institutions that prioritize transparency, institutional obligation, human agency, and innovation. The learning proposed in this model is one that is practical for organizations looking to implement LRMs without diluting accountability.

Furthermore, it serves as a basis for future empirical studies on real data from greater sample sizes and multiple Indian cities as well as comparisons of risk levels across sectors.

7.1 Limitations and Future Research

- Fine grained focus on a single city makes it easier to understand and difficult to make claims about the context for other cities in India that have similar IT systems (Hyderabad, Pune, Chennai, Gurugram, and Mumbai).
- The cross-sectional design cannot establish causation. Investigations of changes in governance maturity over time would benefit from future studies conducted in a longitudinal fashion.
- Self-reported perceptions are used in the model. Future research should incorporate the use of survey data, as well as audit logs, incident reports, model performance measures, and qualitative interviews.
- Going forward, there may be differences in sector/industry needs that can be identified through research by comparing fintech, healthcare tech, manufacturing analytics, and public sector technology vendors for example.

References

[1] Agarwal, Abhinav, and Nene, Manisha J. (2026). A federated architecture for sector-led AI governance: Lessons from India. arXiv. https://arxiv.org/abs/2603.26865

[2] Agarwal, Abhinav, and Nene, Manisha J. (2025). A five-layer framework for AI governance: Integrating regulation, conformity assessment, standards, and assurance. arXiv. https://arxiv.org/abs/2509.11332

[3] Batool, Amna, Zowghi, Didar, and Bano, Muneera. (2024). Responsible AI governance: A systematic literature review. arXiv. https://arxiv.org/abs/2401.10896

[4] Bhalla, Nitika, Brooks, Laurence, and Leach, Tonii. (2024). Ensuring a 'responsible' AI future in India: RRI as an approach for identifying the ethical challenges from an Indian perspective. AI and Ethics, 4, 1409-1422. https://doi.org/10.1007/s43681-023-00370-w

[5] Billah, Md Masum, Hamjaya, Harry Setiawan, Shiralizade, Hakima, Singh, Vandita, and Inam, Rafia. (2025). Large language models' trustworthiness in the light of the EU AI Act: A systematic mapping study. Applied Sciences, 15(14), Article 7640. https://doi.org/10.3390/app15147640

[6] Chappidi, Samhita, Cobbe, Jennifer, Norval, Chris, Mazumder, Anirban, and Singh, Jatinder. (2025). Accountability capture: How record-keeping to support AI transparency and accountability reshapes algorithmic oversight. arXiv. https://arxiv.org/abs/2510.04609

[7] Fabiano, Nicola. (2024). AI Act and large language models: When critical issues and privacy impact require human and ethical oversight. arXiv. https://arxiv.org/abs/2404.00600

[8] George, A. Shaji. (2024). Establishing global AI accountability: Training data transparency, copyright, and misinformation. Partners Universal Innovative Research Publication, 2(3).

[9] Gupta, K. P. (2019). Artificial intelligence for governance in India: Prioritizing the challenges using analytic hierarchy process. International Journal of Recent Technology and Engineering, 8(2S11), 3756-3762.

[10] Harmon, Scott, Wilsmann, Miriam, Joshi, Garima, Ballesteros, Alberto, and Baitinger, Paul. (2024). Decoding India's AI governance strategy and its implications for the U.S.-India bilateral relationship. Indian Public Policy Review, 5(4), 51-82.

[11] Ho-Dac, Marion, and Martinez, Baptiste. (2024). Human oversight of artificial intelligence and technical standardisation. arXiv. https://arxiv.org/abs/2407.17481

[12] Jacob, Joylin, Dorshia, Ancy, and Edwin, Aalan Joe. (2025). Artificial intelligence and legal accountability in India: A socio-legal perspective on emerging regulatory challenges. Indian Journal of Law and Legal Research, 7(5).

Research Article

[13] Janssen, Marijn. (2025). Responsible governance of generative AI: Conceptualizing generative AI as complex adaptive systems. Policy and Society, 44(1), 38-51.
[14] Joshi, Divij. (2024). AI governance in India: Law, policy and political economy. Communication Research and Practice. Advance online publication. https://doi.org/10.1080/22041451.2024.2346428
[15] Kandikatla, Laxmiraju, and Radeljic, Branislav. (2025). AI and human oversight: A risk-based framework for alignment. arXiv. https://arxiv.org/abs/2510.09090
[16] Kathuria, Yatin, Chaki, Nishit Ranjan, Kaur, Manpreet, Kumar, Satish, and Medhavi, Raka. (2024). Responsible AI impact assessment mechanism for India: A robust strategy for effective governance of AI systems in the country. AIP Conference Proceedings, 3220(1), Article 040010. https://doi.org/10.1063/5.0234670
[17] Khatri, Vrinda, and Kewat, Gaurav. (2025). Artificial intelligence in India's legal system: Navigating accountability, liability, and legal voids. International Journal of Human Rights Law Review, 4(2), 509-526.
[18] Konde, Akash. (2026). Artificial intelligence and algorithmic accountability: The need for a legal framework in India. Indian Journal of Legal Review, 6(1), 151-158.
[19] Madaio, Michael A., Stark, Luke, Vaughan, Jennifer Wortman, and Wallach, Hanna. (2020). Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-14). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376445
[20] Meskó, Bertalan, and Topol, Eric J. (2023). The imperative for regulatory oversight of large language models, or generative AI, in healthcare. npj Digital Medicine, 6, Article 120. https://doi.org/10.1038/s41746-023-00873-0
[21] Mittal, Anirudh M., Parthasarathy, Pranav D., and Joshi, Saurabh. (2025). AI ethics education in India: A syllabus-level review of computing courses. arXiv. https://arxiv.org/abs/2509.22329
[22] Mökander, Jakob, Schuett, Jonas, Kirk, Hannah Rose, and Floridi, Luciano. (2024). Auditing large language models: A three-layered approach. AI and Ethics, 4, 1085-1115. https://doi.org/10.1007/s43681-023-00289-2
[23] Novelli, Claudio, Casolari, Federico, Rotolo, Antonino, Taddeo, Mariarosaria, and Floridi, Luciano. (2024). Taking AI risks seriously: A new assessment model for the AI Act. AI & Society. Advance online publication. https://doi.org/10.1007/s00146-023-01723-z
[24] Pahune, Saurabh, Akhtar, Zahid, Mandapati, Venkatesh, and Siddique, Kamran. (2025). The importance of AI data governance in large language models. Big Data and Cognitive Computing, 9(6), Article 147. https://doi.org/10.3390/bdcc9060147
[25] Pandey, Pankaj. (2025). Digital sovereignty and AI: Developing India's national AI stack for strategic autonomy. Procedia Computer Science. Advance online publication.
[26] Papagiannidis, Emmanouil, Mikalef, Patrick, and Conboy, Kieran. (2025). Responsible artificial intelligence governance: A review and research framework. The Journal of Strategic Information Systems. Advance online publication.
[27] Priyanshu, Aniket, Maurya, Yash, and Hong, Zhuo. (2024). AI governance and accountability: An analysis of Anthropic's Claude. arXiv. https://arxiv.org/abs/2407.01557
[28] Puchakayala, P. R. A. (2025). Responsible AI: Ensuring ethical, transparent and accountable artificial intelligence systems. International Journal of Computer Engineering and Technology, 16(2).
[29] Pujari, Tejas. (2025). Ethical and responsible AI: Governance frameworks and policy implications for multi-agent systems. International Journal of Computer Engineering and Technology, 16(2).
[30] Pujari, Tejas. (2025). Explainable AI and governance: Enhancing transparency and policy frameworks through retrieval-augmented generation. International Journal of Computer Engineering and Technology, 16(2).

Research Article

[31] Raji, Inioluwa Deborah, Smart, Andrew, White, Rebecca N., Mitchell, Margaret, Gebru, Timnit, Hutchinson, Ben, Smith-Loud, Jamila, Theron, Daniel, and Barnes, Parker. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (pp. 33-44). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372873

[32] Ramesh, Divya, Kameswaran, Vaishnav, Wang, Ding, and Sambasivan, Nithya. (2022). How platform-user power relations shape algorithmic accountability: A case study of instant loan platforms and financially stressed users in India. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (pp. 1917-1928). Association for Computing Machinery. https://doi.org/10.1145/3531146.3533237

[33] Ribeiro, Diogo, Rocha, Tiago, Pinto, Gustavo, Cartaxo, Bruno, Amaral, Marcelo, Davila, Nilton, and Camargo, Alexandre. (2025). Toward effective AI governance: A review of principles. arXiv. https://arxiv.org/abs/2505.23417

[34] Sambasivan, Nithya, Arnesen, Emily, Hutchinson, Ben, Doshi, Tulsee, and Prabhakaran, Vinodkumar. (2021). Re-imagining algorithmic fairness in India and beyond. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 315-328). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445896

[35] Sharma, Pranjal, Sarma, Anirban, Basrur, Amoha, and Tripathi, Prateek. (2023). AI governance in India: Aspirations and apprehensions. Observer Research Foundation.

[36] Srinivasan, Ramya, and Chander, Ajay. (2021). Biases in AI systems. Communications of the ACM, 64(8), 44-49. https://doi.org/10.1145/3464903