
Managerial Sensemaking in the Era of AI Agents: A Qualitative Research on Leaders Adapting Decision-Making in the Workplace

¹*Hammad Razi

¹Department of Marketing, College of Management Sciences, PAF-KIET University, Pakistan.

Abstract

Artificial intelligence agents have become active partners in the organisational decision-making process, replacing automated tools with autonomous systems that can reason, adapt, and shape the judgement of managers. The current qualitative research paper examines how managers in the fast changing AI environment of Saudi Arabia can understand agentic systems and reorganise decision practices for them. Nine finance, healthcare, operations, marketing, and technology managers were interviewed in semi-structured interviews, which were analysed through reflexive thematic analysis. Five themes were identified, which include evolution of managerial mental models in agentic AI contexts, hybrid agency negotiation: human-agent authority dynamics, autonomous agentic workflows, ethics and governance, reconfiguring managerial labour and value creation with AI-augmentation and sensemaking, trust calibration and organisational culture during AI agent integration. The introduction of AI has brought about the ambiguity of algorithms, which have transformed patterns of confidence, risk perceptions, and delegations, necessitating constant interpretive, ethical, and relational efforts. The research project applies the sensemaking theory to the context of agentic AI and offers real-world knowledge on how to collaborate ethically with AI; however, it requires longitudinal, cross-cultural, and multi-level studies.

Keywords: Artificial Intelligence, AI agents, managerial sensemaking, managerial labour, human-AI interaction, managerial mental models, agentic AI, human AI governance, agency driven labour theory, decision-making.

1 Introduction

Artificial intelligence (AI) agents are becoming the new partners in the organisational decision-making process, beyond the automation of tasks [1]. They can reason, plan, and perform complex tasks. However, managers face the challenge of determining what these agents can or cannot do, distinguishing between hype and practical value, and integrating them into work processes in a meaningful way. The repetitive duties and some predictive information that AI can manage are valid; nevertheless, human control is vital, particularly in high-stakes decisions. A survey of 1,000 enterprise AI developers in 2025 found that 99% are exploring or developing AI agents, marking a transformative development toward agentic adoption [2]. Within this context, leaders must operate in a multi-agent workflow that is becoming increasingly complex, while also ensuring compliance, transparency, and accountability [3]. Adaptation requires strong AI policies and governance systems that preserve human judgment despite the rise in productivity and innovation by the agents. With AI integrated, managerial sensemaking will be essential for

understanding agent capabilities, the trade-off between autonomy and control, and reconciling adoption with organisational goals [4].

Organisational uptake is still in its early stages, despite this momentum. Almost two-thirds of companies remain at the experimental or pilot stage, and 62% are just starting to consider AI agents in their work processes in practice [5]. The initial results are encouraging as 64% of organisations describe better innovation and several report increased cost and revenue performance. Nevertheless, profitability gains are not equalised, and only 39% of them experienced increase in earnings at the enterprise level [5]. There is a sharp difference in orientation between high performers; 80% of them focus on efficiency, whereas those with the highest value focus on growth and innovation, commonly with the help of workflow redesign [6]. About half of these high performers seek transformative change which augurs a more fundamental restructuring of managerial work.

The developments make sensemaking even more important. Managers and employees should observe, derive meaning, and understand AI outputs before incorporating them into

Hammad Razi

Department of Marketing, College of Management Sciences, PAF-KIET University, Pakistan.

Email: hammadrazi@researchergroup.co



©2025 Copyright by the Authors.

Licensed as an open access article using a [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

the decision-making process. This uncertainty is reflected in workforce expectations, with 32% of organisations forecasting workforce downsizing, 43% forecasting no change, and 13% forecasting a workforce increase [5]. The available literature concentrates on the technological drivers, although there is less information about the role of managers in forming interpretive climates and impacting the interpretation of the employees regarding AI agents. This ambiguity presents the necessity to study these dynamics in the context of rapidly developing countries such as Saudi Arabia.

Saudi Arabia offers a unique national perspective in analysing these developments. AI is identified as one of the key drivers in Vision 2030 which supports sophisticated digital infrastructure and widespread cross-sector adoption [7]. The 40 billion AI investment by the Kingdom in 2024 (announced) is one of the largest in the world, and it is estimated that it would contribute USD 135 billion to the GDP by 2030 [7]. The activity of the Digital Government Authority on Large Action Models (LAMs) also sheds light on the hopes for autonomous AI in public and organisational services. Organisations are already modernising their computing infrastructure and experimenting with advanced agents to create a high-momentum environment, where managers must interpret, trust, and operationalise AI outputs at high speed [8].

This research explores these early sensemaking mechanisms and demonstrates how concrete and verifiable AI properties can be utilised to learn tasks and form routines. In contrast, more abstract or opaque properties can facilitate exploratory learning and further negotiation between social and technical structures. In contexts like Saudi Arabia, where rapid AI adoption aligns with Vision 2030 goals, such sociotechnical dynamics shaped by routines, norms, and technological novelty impact organisational adoption and collective sensemaking cycles significantly.

This research aims to identify how Saudi Arabian managers perceive AI agents, negotiate their power and change decision-making behaviour as agents are integrated into the fast-changing organisational processes.

1.1 Theoretical Framework

The present research is based on three interrelated theory lenses: Human AI Governance (HAIG), Agency-Driven Labour Theory (ADLT), and Responsible AI adoption, which aim to explain how managers perceive and adapt AI to organisational decision-making workflows.

- *Human AI Governance (HAIG)*

HAIG provides a comprehensive framework on how to steer AI systems in a way that is reflective of the explicit organisational values, fair, accountable, and transparent [9]. It consists of four pillars, such as value integration, trust and transparency, human empowerment, and strategic alignment, which give managers viable mechanisms to incorporate ethical oversight, enhance collaboration between humans and AI, and make AI adoption work towards organisational objectives and human-centred decision-making.

- *Agency Driven Labour Theory (ADLT)*

ADLT elaborates on how human workers can generate value in the AI-enhanced setting by emphasising on the quality of agency, direction efficacy, and work design [10]. It emphasises the transformed role of managers in facilitating human-AI cooperation, allowing them to make informed decisions and organise workflows in which AI takes on operational roles. ADLT assists leaders in transitioning to AI-based operations, rationalising human contributions, and aligning organisational intent with sustainable and agency-oriented labour relations.

- *Responsible AI Adoption*

Accountable adoption of AI entails the implementation of AI in a manner that is morally, legally, and ethically sound, as well as in line with human and organisational principles [11, 12]. Even though the adoption of AI is increasing at a high pace, the majority of ethics are still based on general theories, which provide minimal guidance. Surveys on 84 standards indicate that there are five principles, including transparency, fairness, non-maleficence, responsibility, and privacy, which are consistently prioritised. As AI systems become increasingly autonomous, organisations must incorporate these principles into their governance systems, ensuring that leaders make ethical, accountable, and strategic AI-powered decisions [13].

Integrated Perspective

All these theories provide a paradigm to explain the interpretive issues, relationships of trust, and behavioural change that emerge when AI agents are introduced into the managerial process of decision-making, as to how AI agents transform processes, thought, feelings, and organisational forms.

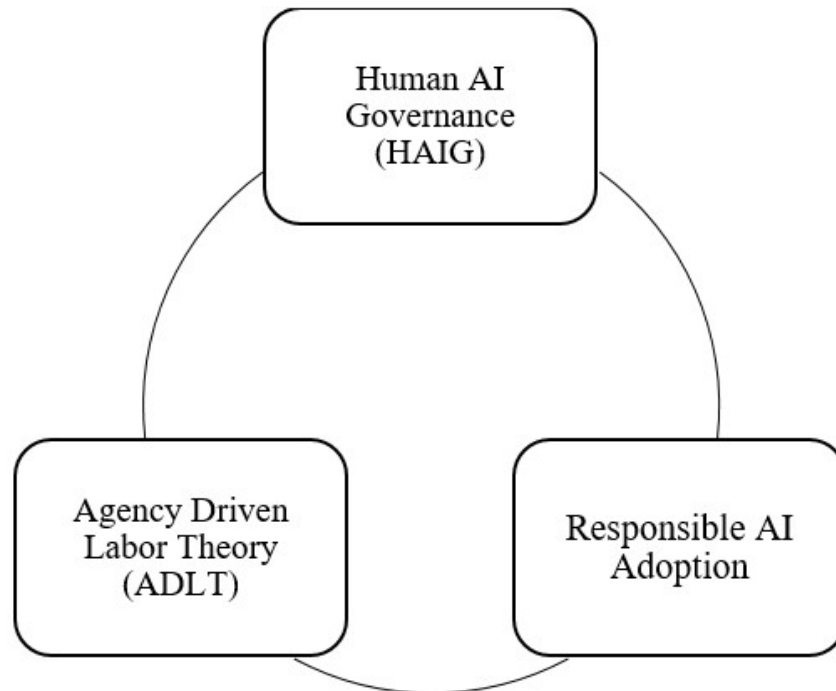


Figure 1. Integrative Theoretical Framework for Managerial Sensemaking in the Age of AI Agents: Interconnections among Human-AI Governance (HAIG), Agency-Driven Labour Theory (ADLT), and Responsible AI Adoption.

2 Literature Review

2.1 Managerial Sensemaking in the Presence of an Algorithmic Ambiguity

As organisations shift from using AI tools to autonomous AI agents, managers are facing a profound ambiguity in algorithms, which is redefining the perception of machine-made decisions [14]. Empirical data extracted from 40 managers revealed that AI-based management systems tend to act as algorithmic bosses, providing unilateral and impersonal feedback, with no room for argument or clarification of context [15]. The result of this reduction of communication to one-way instructions is the emotional dissonance, detachment, and ambiguity since the employees are finding it hard to comprehend the unclear algorithmic directive on how to schedule or prioritise tasks or what is expected of them. Obscurity compels managers to play an interpretive role in decoding inflexible algorithmic results into organisational advice [16]. It indicates that employees with algorithmic control often turn to informal networks to recreate a sense of the shared meaning and regain a sense of agency when the automated instructions are incompatible with the real situation or appear arbitrary.

Complementing these insights, Viktorelius and Larsson (2023) demonstrate that AI technologies produce

several conflicting meanings within an organisation, and managers, frontline workers, and leaders develop various sensemaking frames [17]. It is based on research aimed at examining how professionals sense AI in organisational learning settings, with a specific case of AI use in maritime operations. It shows that the inability to achieve cooperation, learning, and successful adoption of AI is caused by discursive tensions between developers and maritime professionals due to the divergent interpretation of competence, practise, and technology [17].

2.2 Human-AI Collaborative Workflow, Hybrid Authority and Decision-making

Legal scholarship underscores that the concept of hybrid human-AI decision-making is emphasised by the necessity to maintain the human agency, accountability and compliance models [18]. As demonstrated by Enarsson, Enqvist, and Naarttjarvi (2021), hybrid systems place the human in the loop to enhance transparency, legality, and mitigate bias and systemic risks. However, legal judgment by human reviewers is often conflicting in situations where they must make decisions under severe time and resource constraints [18]. In their analysis, they also draw attention to the fact that hybrid decisions are still highly influenced by constitutional rights, administrative law, and power

relations between decision-makers and those affected in practice, but these influences need not be overt; rather, they are context-specific.

Regarding the workflow level, Tarafdar (2025) distinguishes several models of collaboration, such as sequential, parallel, and interactive and illustrates that the trust relies heavily on transparency, predictability, and adherence to user mental models [19]. When systems improve instead of endangering the expertise of workers, the latter are more likely to embrace AI. In addition to this, hybrid intelligence studies indicate that coordinated human-AI cooperation enhances the efficiency and responsiveness of decisions by 10-20% especially when authority is explicitly shared and not completely automated [19].

According to Piridi (2025), there is critical empirical evidence showing that organisations, in which the hand-off boundaries between humans and AI are clearly defined, boast 43% more implementation success and are much more satisfied with the implementation [20]. The case studies on Copilot at Microsoft also reveal that bidirectional hand-off tracking boosts adoption by 37% and productivity by 29%. Clear reasoning in the transition process is necessary, with 72% of knowledge workers citing it as the primary reason why they feel trust [20].

Collectively, these results indicate that hybrid authority is not merely a technical set up but a socio-legal and organisational structure. To ensure successful human-AI collaboration, it is necessary to have clear decision boundaries, handoffs, confidence limits, and powerful oversight mechanisms that retain human judgment while leveraging the analytical capabilities of AI.

2.3 Trust, Ethics and Organisational Culture in the Age of Built-In AI Agents

According to Rezaei et al. (2021), another key factor ensuring the acceptance among employees of embedded AI agents in the decision-making process of organisations is trust [21]. Their results demonstrate that employees will doubt the impartiality and legitimacy of algorithmic outputs when the workload is distributed or performance rated algorithmically with increased levels of opacity and reduced explainability and perceived surveillance. The research about the Iranian e-retail business enterprises determined major challenges that affect AI-based decision-making in knowledge systems. It was done with CFA on 198 valid answers and it demonstrated that the issues of the highest importance were privacy and data protection, bias and fairness, and transparency and explainability,

with strong I coefficients (up to 0.88) and high predictive validity in decision-making situations [21].

Likewise, Singh (2021) suggests that tensions of ethical concerns are escalated as AI systems induce organisational biases or alter the accepted norms of accountability, which leads to the uncertainty of the decision makers in hybrid human-AI workflows [22]. The empirical research by AI Samman (2024) also shows that AI transforms organisational culture, changing the communication patterns and expectations regarding transparency [23]. The respondents in his research emphasised that leadership transparency, effective communication, and regular value coherence are key to maintaining trust in AI-capable environments. Murire (2024) continues that despite its ability to enhance cultures of innovation and lifelong learning, ethical matters related to AI, including bias, privacy, and unequal power balance, are still touched upon by employees, which is why ethical governance and cultural preparedness are both essential success factors [24]. Taken together, these studies show that trust and ethics are deeply intertwined with organisational culture, and successful adoption of AI agents depends on creating environments where transparency, accountability, and shared values guide human-AI interaction.

All of this indicates that trust and ethics are closely connected to organisational culture, and the successful implementation of AI agents requires the establishment of an environment where transparency, accountability, and shared values can regulate human-AI interaction.

Although much has been researched on the use of AI, there is limited empirical evidence that managers consider AI outcomes, determine power of authority, and assess levels of belief in the actual organisational processes. This research addresses these gaps by qualitatively exploring managerial sensemaking and adaptation in cases involving the use of AI agents.

3 Methodology

The research employed an interpretivist methodology, which is qualitative, and examined how managers in Saudi Arabia perceive and accommodate AI agents in practical organisational contexts. Since the process of AI-related decision-making implies subjective interpretation, emotional reactions, and cognitive adaptation, the qualitative method was the most appropriate. In contrast to quantitative designs, which identify patterns of use, this approach reveals the underlying sensemaking, negotiations, and experiential senses of meaning that occur

as AI is integrated into managerial processes, noting how managers form and apply insights generated by AI.

3.1 Sampling

Managers in Saudi Arabia with long-term, practical experience in using AI for decision-making were recruited using purposive sampling. Individuals were recruited from areas including finance, operations, marketing, healthcare, and technology, where AI aids in prediction, risk assessment, and workflow automation. To be included, they needed to have at least two years of experience in a managerial role and regularly engage with AI systems, such as analytics dashboards, generative assistants, predictive engines, or automated screening tools. Individuals who had minimal or incidental use of AI were eliminated. The sample size of 7-10 managers was deemed adequate to provide depth, while also achieving thematic saturation.

3.2 Data Collection

Data were analysed using Braun and Clarke's six-phase reflexive thematic analysis, which aligns with the study's interpretivist focus by prioritising meaning-making over mechanical coding [25]. Analysis began with repeated transcript readings for deep familiarisation, followed by systematic coding of patterns related to interpretation, trust, emotional response, and behavioural adaptation. Preliminary themes were generated and iteratively refined against the full dataset to ensure coherence and representativeness, then clearly defined and named to form the study's interpretive narrative. NVivo software supported efficient organisation and retrieval of codes, strengthening transparency and analytical rigour. Coding remained reflexive and iterative throughout, acknowledging the researcher's active role in constructing insights.

3.3 Data Analysis

The data were interpreted through a six-phase reflexive thematic analysis, suggested by Braun and Clarke, which is appropriate in the present research as it is an interpretivist approach, meaning that it does not require scholarly, mechanical coding of data [25]. The analysis process commenced by reading transcripts multiple times to become well-acquainted with them, followed by a systematic analysis of themes associated with interpretation, trust, emotional response, and behavioural adaptation. Themes were initially created and repeatedly

adjusted to the total data to ensure consistency and representativeness, and finally articulated and described to create the interpretation story of the research. Transparency and analytical rigour were facilitated by NVivo software, which enabled effective organisation and retrieval of codes, thereby enhancing efficiency and facilitating the analysis. Coding was reflexive and iterative; the construction of knowledge involved the active role of the researcher.

4 Findings

A reflexive thematic analysis was conducted, and transcripts were manually coded. Nine managers participated in this research and will be referred to as T1 through T9. Every interview was recorded with the participant's informed consent and transcribed verbatim by the researcher. This led to the emergence of five central themes: Evolution of managerial mental models in agentic AI contexts, hybrid agency negotiation: human-agent authority dynamics, autonomous agentic workflows, ethics and governance, reconfiguring managerial labour and value creation with AI-augmentation and sensemaking, trust calibration and organisational culture during AI agent integration.

In the current research (2023-2025), the researchers outline the notion of managerial sensemaking with AI agents, focusing on the following critical themes: changing managerial mental models, hybrid agency negotiation, autonomous workflow ethics and governance, reconfiguring managerial labour, and calibrating trust within organisational culture. Although these themes are connected to frameworks such as HAIG and Responsible AI Adoption, little empirical research has been conducted, especially on the dynamics of authority in real-life scenarios, over time, and subtle sensemaking, which are the main points of focus in future research on AI-enhanced managerial scenarios.

Theme 1: Evolution of Managerial Mental Models in Agentic AI Contexts

Managerial mental models in agentic AI situations originate from the way managers continually reframe AI-generated insights to integrate algorithmic logic into established cognitive and decision-making frameworks. Managers have been proactive in redesigning their mind framework to embrace AI insights. Instead of thinking of AI as a neutral entity, respondents described it as a second analyst, risk filter, or pattern-detection engine, rather than

Table 1. Novel Research Themes in Managerial Sensemaking and Decision-Making with AI Agents (2023-2025): Gaps, Framework Relevance, and Contributions

Theme	Sub-Themes / Analytical Focus	Focus / Description	Relevance to Frameworks	Research Gap / Novelty
Evolution of Managerial Mental Models in Agentic AI Contexts	Cognitive adaptation to AI outputs Interpretive frameworks for AI-generated insights Scenario experimentation and pattern recognition Sensemaking under uncertainty	Examines how managers' cognitive schemas adapt as they collaborate with autonomous AI agents.	HAIG, Responsible AI Adoption - highlights governance and ethical alignment in decision-making.	Limited empirical research on evolving mental models under agentic AI; need to study interpretive adaptation.
Hybrid Agency Negotiation: Human-Agent Authority Dynamics	Decision authority negotiation between humans and AI Role redefinition and delegation Verification and override behaviours Power dynamics in hybrid decision-making	Explores negotiation of decision authority between managers and AI agents, including role redefinition.	ADLT, HAIG - focuses on human agency, direction effectiveness, and governance of AI decisions.	Few studies analyse real-world negotiation of authority in hybrid human-AI decision systems.
Autonomous agentic workflows, Ethics and Governance	Trust and transparency in autonomous AI decisions- Accountability and professional responsibility Ethical dilemmas in AI-driven workflows Compliance with organisational or regulatory policies	Investigate how ethical and accountability frameworks adapt to AI agents acting autonomously in workflows.	HAIG, Responsible AI Adoption, emphasises trust, transparency, and ethical integration.	Existing frameworks are conceptual; empirical models for autonomous workflows are scarce.
Reconfiguring Managerial Labour and Value Creation with AI-Augmentation	Role transformation and delegation patterns Hybrid human-AI decision-making models Value creation through AI augmentation Cognitive load and efficiency optimisation	Studies how AI agents shift managerial roles, judgment, and value creation.	ADLT - addresses agency quality, framework design, and direction effectiveness in labour dynamics.	Lack of longitudinal research on how managers adapt roles and create value with agentic AI.
Sensemaking, Trust Calibration and Organisational Culture during AI Agent Integration	Trust calibration in human-AI interactions Emotional responses and confidence in AI-driven decisions- Organisational culture and AI adoption norms- Skills, training, and resource constraints affecting AI integration	Focuses on how managers interpret AI outputs, calibrate trust, and cultivate a supportive organisational culture.	HAIG, Responsible AI Adoption - links governance, accountability, and responsible adoption to managerial cognition.	Current research emphasises trust, but nuanced sensemaking and cultural adaptation remain underexplored.

situating AI in cognitive and organisational frameworks. T1 described: "I consider AI as an additional analyst in my work team. It identifies patterns that I might have missed, but I will

need to confirm them. It is not taking my place; it is about having another angle on the same problem." The managers used specific interpretive practices in cross-referencing AI results against past patterns, departmental

standards, and their own experience. T4 noted:
“The sharp forecast change shown by AI prompts me to be the first to check what happened in the previous quarter or whether something out of the ordinary has occurred recently. AI is intelligent, but not all situations are the same.”

One of the strategies was scenario experimentation. T7 reflected:

“When I change something and the model wigs out and goes backwards, I know it is delicate. I will not ignore an important trend, in case it continues.”

This theme highlights patterns of constant loops of observing, doubting, and confirming AI results, which is consistent with sensemaking in the presence of uncertainty. Managers combine AI information and experience with organisational knowledge, emphasising the presence of interpretive adaptation as an essential attribute of agentic AI cooperation.

Theme 2: Hybrid Agency Negotiation: Human-Agent Authority Dynamics

The concept that captures such roles is hybrid agency negotiation, which involves a dynamic balance of redistributing decision-making authority between managers and AI agents, as well as renegotiating roles, oversight, and control within the collaborative decision-making process. Human-AI collaboration was based on authority negotiation. According to the reports of managers, situational delegation of decision-making occurred, and the balance between AI recommendations and human judgment was modified by the complexity of the tasks, their reliability, and ethical considerations.

T3 explained:

“When it comes to routine forecasting, I leave AI at the forefront. However, when making strategic investments, I disregard advice, even if the confidence of the AI is high.”

To ensure accountability and transparency, such negotiations were also recorded by participants. T5 stated:

“We document our reasons for accepting or denying AI recommendations. This is beneficial when discussing the results or justifying decisions before stakeholders.”

Shifts in role perception occurred among managers. T8 noted:

“The experience of sharing decision space with AI changed my perception of power. Control is less about oversight, strategy, and checks on ethics.”

This theme highlights dynamic role negotiation and a focus on hybrid agency, where human power is realigned

in real time to combine AI capabilities with responsible integration.

Theme 3: Autonomous agentic workflows, Ethics and Governance

Ethics and regulation of autonomous agentic workflow: The issue of how managers can maintain accountability, transparency, and ethical regulation as AI agents gain greater autonomy in organisational processes. When AI went on its own, managers emphasised the importance of ethical control and responsibility. There was active construction of trust in AI by validation, openness and compliance with organisational policies.

T2 observed:

“My question before acting on AI outputs is: is it ethical, compliant and defensible? AI does not perceive consequences in the same way that human beings do.”

Responsibility was a fundamental issue of interest. T6 added:

“Although AI may provide a recommendation, it is my responsibility to achieve the desired results. I ensure that there is a reason behind every choice, and one of them is related to AI.”

The relationship between responsible AI adoption and organisational ethics can be illustrated by the way managers use policies to strike a balance between autonomy and oversight.

Theme 4: Reconfiguring Managerial Labour and Value Creation with AI-Augmentation

Reconfiguring Managerial Labour and Value Creation under AI-Augmentation examines how AI transforms managerial work, reassigns thought, and enables new types of strategic and organisational value. The augmentation of AI recalculated the role of managers, delegation patterns, and value creation. Managers stated that there was less cognitive load on routine activities, thereby allowing for strategic attention to be focused on other tasks.

T1 noted:

“AI performs regular analytics, and I specialise in communication, interpretation and strategy. It is a change from doing to directing.”

Value creation was not just about operational efficiency. T4 reflected:

“With AI, we can identify opportunities that we would not have noticed otherwise, such as a new market opportunity or a fragile risk profile. That brings physical company

value or worth.”

Prioritisation and mental load were also proactively addressed. T9 explained:

“I do not stop making decisions, but AI eliminates noise. I spend more time on high-impact choices and less on routine tasks.”

This theme provides for a change in roles and hybrid decision-making, demonstrating how managers utilise AI to optimise labour, enhance judgment, and develop strategic opportunities.

Theme 5: Sensemaking, Trust Calibration and Organisational Culture during AI Agent Integration

The sensemaking, trust calibration, and organisational culture in integrating AI agents explain how managers interfere with the outputs of AI and regulate the level of their trust in algorithmic decisions, addressing cultural norms to help their collaborations with AI agents remain useful. The concept of trust calibration proved to be a dynamic process, shaped by the reliability of AI, the experience of working in the sphere, and organisational culture. Feelings influenced the degree of confidence in AI-mediated decisions.

T3 remarked:

“I put my trust into AI when it keeps on moving in the right direction of established trends, yet a single mistake can undermine my faith.”

The use of AI was directed by organisational culture and norms. T8 stated:

“Guidelines and training enabled me and my team to use AI with comfort. Technology is as value-neutral as culture.”

Simple resource and skill constraints are the primary factors mentioned by the participants, as they argue that implementing AI properly requires both technical preparedness and cultural compatibility.

In all the thematic areas, AI has been observed as a proactive partner, encouraging managers to participate in the interpretative processes, negotiating authority, addressing ethics, and reorganising labour, while calibrating trust. Metaphoric conceptualisation carries cognitive adaptation and dynamic authority negotiation, depicting hybrid agency. Role transformation enables the creation of strategic value, and cultural integration, facilitated by trust calibration, makes human-AI collaboration effective.

5 Discussion

The findings indicate that Saudi managers are undergoing a significant cognitive transformation, to the

extent that they consider AI agents as part of their daily and strategic decision-making. In line with Scarbrough et al. (2025), who note that AI generates numerous and at times conflicting meanings in professional circles, managers in this research term AI as an additional analyst, risk filter, or pattern detector, with mental model rebuilding evidently occurring. This is attributed to the argument by Engstromb et al. (2024) that AI is driving sensemaking away from episodic interpretation to continuous reflection, as managers are faced with opaque and probabilistic outputs. This interpretive labour is intensified, especially in the case of Saudi Arabia, where AI implementation is supported by the state investment and fast digitalisation. Triangulating AI output with previous trends, testing scenario variations, and verifying anomalies are activities in which managers are involved, indicating a shift in cognitive processing toward a hybrid mode of operation. In general, the development of mental models demonstrates how AI agents tend to influence the reasoning of managers as well as their approaches to uncertainty and strategic foresight.

The negotiation of decision-making authority between human and AI agents has become one of the key characteristics of managerial work in Saudi organisations. Based on Emarsson et al. (2022), who observe that under hybrid systems, human evaluators must strike a balance between legal responsibility and AI-based suggestions, participants mentioned that they have taken an active role in deciding when to revert to AI or disregard it. This can be discussed, as noted by Tarafdar (2025), who states that transparency and predictability have a significant impact on the acceptance of hybrid workflows. In this research, managers regularly recorded override decisions to ensure ethical defensibility and organisational legitimacy, and to confirm the applicability of human judgement in situations of high stakes or uncertainty. Another change in managerial functions was also observed by Saudi managers, who mentioned directive authority and supervisory oversight as characteristics of redefined managerial functions, which aligns with the focus of ADLT on preserving agency. Consistent with Piridi's (2025) findings regarding hand-off design, organisations with well-structured AI-human boundaries were better coordinated and trusted to a greater extent. Therefore, the hybrid agency is not only technical, but also a negotiated social relationship, based on the norms of accountability and the development of managerial identity.

Saudi managers indicated that they were increasingly sensitive to the ethical implications of delegating aspects

of decisions to AI agents, as expected in the global literature on moral concerns, which identifies this as a key obstacle to the adoption of AI. Rezaei et al. (2024) demonstrate that the flaws in the algorithmic decision-making process result in low levels of trust, which is also evident in the participants' unwillingness to accept the algorithm's messages and insistence on the ability to verify the information and make it defensible. Managers argued that responsibility is inherently human, especially in controlled markets like healthcare and finance. This reflects the idea that hybrid systems may cause blurred sets of responsibility, as proposed by Singh and Chandra (2024), who suggest that culturally and organisationally aligned structures of governance are required. Manager expectations in Saudi Arabia are another area where the concept of responsible AI is institutionalised at the national level, as the participants cited the policies and ethics of organisations as points of reference in the decision-making process. In general, managers perceived the impact of governance not as a compliance activity, but as a vital process of ensuring legitimacy, fairness, and trust as autonomous AI workflows continue to grow.

In line with the Agency-Driven Labour Theory (ADLT), the results indicate that AI agents are transforming managerial labour, diminishing routine analytical requirements and reallocating efforts to strategic interpretation, coordination, and communication with stakeholders. The subjects reported efficiency improvements and reduced cognitive load, which is consistent with the wider industry finding by Tarafdar (2025) that a 10-20% improvement is observed when hybrid intelligence systems are designed appropriately. Saudi managers have noted the usefulness of AI, especially in revealing underlying trends, predicting risks, and identifying new business opportunities, with the implication that AI can be used to add financial intelligence to organisations. This role reparation proved most apparent in areas that are accelerating rapidly in the digital sector, where AI allowed managers to focus on the decisions that have the most impact. Nevertheless, the importance of human expertise remained unaffected by this increment; it rather enhanced the interpretative and integrative elements of managerial labour. The results have demonstrated that value creation needs to become interdependent on the capacity of managers to merge AI understandings with contextual judgement, cultural knowledge, and organisational goals.

The research findings indicate that trust calibration is a dynamic and experience-based phenomenon that is

determined by reliability, organisational culture and emotional reactions. In line with Rezaei et al. (2024), managers indicated that trust is a slow process that is built on performance but it can be undermined by one unforeseen mistake. The cultural aspects were critical; organisations that excelled at communication processes, transparency standards, and organised AI training integrated without hiccups according to the results by Al Samman (2024), showing that AI reconstructs and is reconstructed by the cultural expectations. The Saudi managers emphasise that the adoption of AI should be in line with organisational attitudes, leadership practices, and employee willingness. It can be argued that Murire (2024) claims that AI can reinforce innovation cultures only when ethical and psychological issues are considered. The relational aspect of trust was noted through the emotional responses, including the calming effect of AI confirming the correctness of choices and the uncomfortable feeling when the results failed to align with the intuition. In general, the effective introduction of AI into Saudi working environments requires not only technical excellence but also cultural consistency, normative clarity, and organisational dedication to open and accountable introduction.

6 Conclusion

In this research, the application of AI as agents in the organisation is depicted as continuous sensemaking by managers in Saudi Arabia. Instead of accepting what the algorithms turn out to be, the managers engage in interpreting, verifying, and putting them into context, reflecting the ambiguity and fluidity of agentic AI. Control and authority over decision-making shift to the negotiated level, and human control, ethical justification, and logical argumentation remain in focus. Reliability, transparency, organisational norms, and emotional reactions between confidence and warning are all sources of dynamic trust. AI equally restructures the managerial workforce, recycling other routine analytical work and increasing strategy analysis, judgement and coordination.

7 Strengths and Limitations

One of the strengths of this research is the author's concentration on the rapidly developing AI environment of Saudi Arabia, which is governed by a variety of managerial opinions with robust reflexive thematic analysis. Member reflections, audit trails, and peer debriefing were used to enhance credibility, and we utilised rich contextual detail to improve transferability. Nonetheless, the data is

qualitative and self-reported, which restricts the possibility of generalisation, and the results can be summarised as the data of a single moment within a rapidly evolving setting. Interpretation may also be distorted as a result of cultural specificity and lack of team-based or immediate observations.

8 Future Research

Future studies should be implemented using a longitudinal design that follows the dynamics of trust calibration, authority negotiation, and sensemaking as AI becomes more developed in organisations. Cross-cultural studies may perhaps explain the influence of national values and organisational norms on the dynamics between humans and AI. Research on explainability, the mechanisms of escalation, and hybrid decision models, as developed through experiments, would provide practical advice when designing a system. Multilevel research integrating teams, organisational culture, and sectoral diversity would help provide a deeper understanding of how AI agents transform the work of managers in various contexts.

Acknowledgements

Funding

No funding required for this study.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

Data will be available upon request.

Ethical Considerations

Ethical approval was obtained, informed consent was secured, and all participant data were handled confidentially and in full accordance with the principles of qualitative research ethics.

References

Booyse, D. and C.B. Scheepers, Barriers to adopting automated organisational decision-making through the use of artificial intelligence. *Management Research Review*, 2024. 47(1): p. 64-85.

Ivan Belcic, C.S. AI agents in 2025: Expectations vs. reality. 2025; Available from: <https://www.ibm.com/think/insights/ai-agents-2025-expectations-vs-reality>.

Mequanenit, A.M., et al., A multi-agent deep reinforcement learning system for governmental interoperability. *Applied Sciences*, 2025. 15(6): p. 3146.

Neiroukh, S., O.L. Emeagwali, and H.Y. Aljuhmani, Artificial intelligence capability and organizational performance: unraveling the mediating mechanisms of decision-making processes. *Management Decision*, 2024.

Hill, Z. McKinsey: Most firms stuck in AI pilot mode despite global surge. 2025; Available from: <https://lionherald.com/mckinsey-most-firms-stuck-in-ai-pilot-mode-despite-global-surge/>.

Company, M. The state of AI in 2025: Agents, innovation, and transformation | Survey. 2025; Available from: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>.

Authrurity, D.G., Generative AI in Digital Government. 2025.

Krishnan, N., Ai agents: Evolution, architecture, and real-world applications. arXiv preprint arXiv:2503.12687, 2025.

Engin, Z., Human-AI Governance (HAIG): A Trust-Utility Approach. arXiv preprint arXiv:2505.01651, 2025.

Ganuthula, V.R.R., Agency-Driven Labor Theory: A Framework for Understanding Human Work in the AI Age. arXiv preprint arXiv:2501.01448, 2024.

Hernández, E.G., Towards an ethical and inclusive implementation of artificial intelligence in organizations: a multidimensional framework. arXiv preprint arXiv:2405.01697, 2024.

Papagiannidis, E., Responsible AI governance in practice: The strategic impact of responsible AI governance on business value and competitiveness. 2024.

Hossain, S., M. Fernando, and S. Akter, The influence of artificial intelligence-driven capabilities on responsible leadership: A future research agenda. *Journal of Management & Organization*, 2025: p. 1-25.

Ahmed, M., Artificial Intelligence, Nuclear Warfare and Towards Laws: A Study on Limitations of Existing International Laws and Treaties. 2025.

Ruslan, A. and R. Saleh, The Algorithmic Boss: Communication Challenges in Organizations Using AI-Driven Management Systems. *Fast in Social Sciences*, 2025. 1(1): p. 90-100.

Lippert, I., Algorithmic Management in Different Work Contexts: Implications for Workers and Managers. 2025.

Viktorelius, M. and S. Larsson, Talking about machines: discursive sensemaking and workplace learning in the age of AI. *Studies in Continuing Education*, 2025: p. 1-19.

Enarsson, T., L. Enqvist, and M. Naarttijärvi, Approaching the human in the loop—legal perspectives on hybrid human/algorithmic decision-making in three contexts. *Information & Communications Technology Law*, 2022. 31(1): p. 123-153.

Tarafdar, R., Human-AI Collaboration in Workflow Optimization: A Framework for Hybrid Decision Systems in Automation-Heavy Industries. 2025.

Piridi, S., Designing Human–AI Hand-Offs: Copilot in Hybrid Workflows. *Journal of Computer Science and Technology Studies*, 2025. 7(5): p. 605-611.

Rezaei, M., M. Pironti, and R. Quaglia, AI in knowledge sharing, which ethical challenges are raised in decision-making processes for organisations? *Management Decision*, 2024.

Singh, D. and S. Chandra, Mitigating Uncertainty and Enhancing Trust in AI: Harmonizing Human-Like,

System-Like Features with Innovative Organizational Culture. 2024.

Al Samman, A.M. The Use of AI in Fostering and Embracing Organizational Culture. in *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*. 2024. IEEE.

Murire, O.T., Artificial intelligence and its role in shaping organizational work practices and culture. *Administrative Sciences*, 2024. 14(12): p. 316.

Ahmed, S.K., et al., Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health*, 2025. 6: p. 100198.

Supplementary Tables

Table 2. Interview Process to Research Managerial Sensemaking, Decision-Making, and Trust Dynamics in AI-Added Organisational settings (2023-2025): Themes, Sub-Themes, and Guide questions

Main Themes	Interview Questions
Evolution of Managerial Mental Models in Agentic AI Contexts	1. How do you conceptualise and integrate AI-generated insights into your existing cognitive frameworks? 2. Can you describe a scenario where AI insights challenged your interpretation or decision-making? How did you adapt?
Hybrid Agency Negotiation: Human-Agent Authority Dynamics	1. In what specific situations do you explicitly negotiate decision authority between you and AI agents, and how do you document or reflect on that negotiation? 2. How do you determine when to override an AI recommendation, considering not only accuracy but also organisational, ethical, or strategic implications? 3. How has your sense of managerial authority or role changed as a result of sharing decision space with AI agents, and how do you manage or track that change?
Autonomous agentic workflows, Ethics and Governance	1. How do you assess the ethical implications of AI-generated decisions in your workflow? 2. How do you manage accountability when AI contributes to decision outcomes?
Reconfiguring Managerial Labour and Value Creation with AI-Augmentation	1. How has your role description, workload, or time allocation changed as AI augmentation was introduced, and how do you quantify or qualify that shift? 2. In what ways has the hybrid human-AI decision-making model changed how value is created in your organisation (e.g., new metrics, outcome types, stakeholders)? 3. How has your cognitive load, prioritisation strategy, or decision-making efficiency been altered by AI, and how do you track whether these changes are beneficial or problematic?
Sensemaking, Trust Calibration and Organisational Culture during AI Agent Integration	1. How do you calibrate trust in AI outputs, and what factors influence this trust? 2. How do organisational culture, norms, or policies influence your use of AI in decision-making?