ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

Agentic AI and Organizational Decision-Making: A Conceptual Framework for Human–AI Collaboration in Digital Transformation

Asher Ojuok[1]

[1]Doctoral Candidate, School of Social Sciences and Technology, Technical University of Kenya

Publication Date: 2026/10/24

Abstract: Organizations are now implementing AI technology that goes beyond prediction or recommendation. Agentic systems can pursue a goal, plan steps, call software tools and adapt with little prompting, which unsettles long-standing assumptions about who decides and who answers for the outcome in digital transformation. In this study, a conceptual framework is designed in relation to collaborative work between humans and artificial intelligence in organization decision making in agentic circumstances. With the use of a purposive narrative review of peer reviewed literature and policy documents published from 2019 to 2026, the issue is contextualized within the current global trends, African policy regarding the continent and Kenya's digital economy. The framework will be designed through the process of iterative synthesis based on three decision attributes (stake, ambiguity, reversibility) and the amount of autonomy assigned to the agent in four types of collaboration; delegate, supervise, co-decide, and reserve. Six testable propositions follow. The paper argues that organizations face a design problem of how to share decisions, not a choice between human and machine decisions, and that in Kenya the design must also respect legal limits on solely automated decisions under section 35 of the Data Protection Act, 2019. Implications for managers and policymakers are outlined, and an agenda for empirical testing is proposed. As a conceptual contribution, the paper reports no new data, and its propositions remain to be tested.

Keywords: *Agentic AI; Organizational Decision-Making; Human–AI Collaboration; Digital Transformation; Kenya.*

How to Cite: Asher Ojuok (2026) Agentic AI and Organizational Decision-Making: A Conceptual Framework for Human–AI Collaboration in Digital Transformation. *International Journal of Innovative Science and Research Technology*, 11(10), 1559-1567. https://doi.org/10.38124/ijisrt/26oct568

I. INTRODUCTION

Prediction has been at the heart of the AI discussion within firms for much of the past decade: a model assessed a loan applicant, identified a problem with a machine or predicted future demand, and a human took it from there. The decision was left to the human, and the technology remained secondary. This pattern is being turned on its head. There is a new type of technology known as agentic AI that can accept an objective stated in natural language, plan the process needed to achieve it, make use of external tools like databases, payment gateways and email, assess its own progress, and try again if things go awry (Yao et al., 2023; Acharya et al., 2025).

This matters for digital transformation in a practical way. However, digital transformation has never been about just the software. According to Vial (2019), digital transformation is a process whereby digital technologies cause strategic reactions and structural change, and according to Verhoef et al. (2021), it causes changes in business models, customer experience, and operations altogether. Agentic AI introduces yet another element into this equation since it changes the locus of action. Earlier systems automated tasks that managers had already defined. Agentic systems can interpret a goal and choose the tasks themselves. When that happens, questions that used to be settled by job descriptions and approval chains come back open. Who authorised this action? Who is responsible if it harms a customer? How does a manager keep real oversight over something that works faster than she can read?

The academic literature has useful pieces of an answer but not a complete one. Work on organizational decision structures shows that AI changes how decisions are centralised and where information flows (Shrestha et al., 2019). Work on the automation–augmentation paradox shows that organizations that lean only on automation or only on augmentation both run into trouble (Raisch & Krakowski, 2021). Baird and Maruping (2021) offer one of the few theoretical treatments of delegation to and from agentic information systems. Yet most of this research was written before the recent wave of agent systems, and much of it treats AI as a single kind of thing. Practitioners now face a more concrete problem. They need to know which decisions to hand to an agent, which to keep, and how to organize the space in between.

This paper takes up that problem. Its aim is to develop a conceptual framework that explains how organizations can

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

structure human–AI collaboration in decision-making when AI is agentic. It asks three questions. First, how does agentic AI change the nature of organizational decisions? Second, what is the basis for the allocation of work between people and agents for a particular decision? Third, what conditions are required to ensure that allocation works, in particular, in environments like Kenya, where the conditions are different from those in the literature?

The paper makes three contributions. It clarifies what is distinctive about agentic AI for decision theory by separating levels of agent autonomy from types of decisions. It proposes a framework with four collaboration modes and four organizational enablers, and it states six propositions that can be tested. And it brings an African and specifically Kenyan perspective into a conversation that has been led mostly by research from North America, Europe and East Asia. The paper is conceptual. It does not report new empirical data, and its claims should be read as propositions to be examined, not as findings.

The paper is organized in accordance with the organization of the journal. The literature review puts the problem into global, continental, and Kenyan perspectives and provides the theoretical background. The methodology describes the development of the theoretical framework. The results include the theoretical framework and propositions. The discussion highlights the implications of the research in terms of theory, management, and policy.

II. REVIEW OF LITERATURE

The literature reviewed here covers peer-reviewed work and official policy documents published between 2019 and 2026. Table 1 summarises the main research streams and the gap each leaves. The sections that follow move from the global picture to the African continental perspective and then to Kenya, before setting out the conceptual foundations of the framework.

Table 1 Summary of Research Streams Reviewed

Research stream	Key sources	Main insight	Gap for this study
Digital transformation	Vial (2019); Verhoef et al. (2021)	Digital technologies reshape strategy, structures and business models together	Little on technology that acts on goals it interprets
Decision structures and AI	Shrestha et al. (2019); Agrawal et al. (2019)	Suitable structure depends on decision features; AI lowers the cost of prediction and raises the value of judgment	Treats AI mainly as adviser, not actor
Agentic systems and delegation	Baird & Maruping (2021); Murray et al. (2021); Yao et al. (2023); Acharya et al. (2025); Sapkota et al. (2025)	Delegation to agentic systems and conjoined agency change how work is done	Few guides on matching collaboration mode to decision type
Human–AI collaboration and trust	Dellermann et al. (2019); Logg et al. (2019); Glikson & Woolley (2020); Lebovitz et al. (2022); Fügener et al. (2022)	Trust and reliance are uneven; joint performance can fall below AI alone	Oversight design linked to decision stakes is thin
Governance and ethics	Jobin et al. (2019); Selbst et al. (2019); Green (2022); European Union (2024)	Principles converge and rules are arriving, but human oversight often works poorly in practice	What kind of human involvement fits which decision
Africa and Kenya	African Union (2020, 2024); Gwagwa et al. (2020); Birhane (2020); Eke et al. (2023); Government of Kenya documents (2019–2025)	Local data, skills, infrastructure and law shape AI use	Organization-level decision design in African settings

➢ *The Global Picture*

Since the late 2010s, interest in the application of AI at an organizational level has been increasing gradually. In the early reviews like those by Duan et al. (2019), on decision-making in the age of big data, and Dwivedi et al. (2021) on multidisciplinary perspectives on artificial intelligence, there has been indication of the mismatch between the capabilities of the technology and the ability of the organization to handle it. With the development of foundation models and their rapid deployment following 2022, the mismatch has grown wider (Maslej et al., 2025). However, what has changed most recently is not only the quality of models but how the models are deployed. Rather than being capable of answering one question at a time, the models are packaged into software that can help them plan, manipulate objects, and perform actions step-by-step. This distinction is well made by Sapkota et al. (2025) in distinguishing single-purpose AI agents from agentic systems comprising a number of agents.

Adoption data suggest that organizations are moving, though unevenly. Survey work by McKinsey (2025) reports that a growing share of firms are experimenting with AI agents, while a smaller group has begun to scale them beyond pilots. The picture is also one of caution. Gartner (2025) predicted that more than 40 percent of agentic AI projects would be cancelled by the end of 2027 because of rising costs, unclear business value or weak risk controls. I read these two signals together. Enthusiasm is real, but many organizations do not yet know how to place an agent inside an existing decision process.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

Regulation is also taking shape. However, the EU's Artificial Intelligence Act of 2024 mandates that high-risk AI systems be developed in a manner that allows natural persons to be able to effectively control them (European Union, 2024). However, while this may appear straightforward, research has cast doubt on the ability of human oversight to achieve regulatory objectives. Green (2022) believes that human in the loop policies tend to assume falsely that people will be able to detect any mistakes made by algorithms and can even place blame on the individual overseeing the process. Jobin et al. (2019) discovered that even though ethical guidelines worldwide tend to converge on certain values like transparency, fairness, and accountability, they differ in their interpretation and implementation of these values. Therefore, the message from around the world is that guidelines exist, and regulation is coming.

➢ *The African Continental Perspective*

African nations come to this debate at a point quite distinct from those already mentioned. Indeed, the Digital Transformation Strategy for Africa 2020–2030 by the African Union envisaged an "integrated digital market and economy where connectivity, digital skills, and trust were among shared concerns" (African Union, 2020). Developing further from these basics, the 2024 Continental Artificial Intelligence Strategy by the same organization stressed the significance of AI adoption in the context of the development of Africa (African Union, 2024).

There are researchers who point out both the benefits and possible risks. For instance, Gwagwa et al. (2020) consider the adoption of AI in the African continent and note the positive effects in such spheres as agriculture, healthcare, and financial sector while stating that there is no local data, poor infrastructure and lacking regulation in this area. The compilation of perspectives from Africa on the topic of responsible AI has been conducted by Eke et al. (2023), which shows that ethical frameworks of other countries cannot be used directly. In addition, Birhane (2020) talks about algorithmic colonization, implying the extraction of data from African citizens and designing of systems based on these data without any regard to local context. This implies that for African organizations, the adoption of agentic AI should include some more aspects than just technical and managerial ones.

Such worries are relevant to decision-making. An agent that is primarily trained using information coming from other areas may interpret wrong the local behavior regarding credit, language or informal markets. The human-in-the-loop approach may be useful, provided that the people who make up the loop have the necessary training and authority. Continental policy documents recognise the need for skills and governance. Less has been said about how individual organizations should decide which decisions to share with an agent, and that is where the framework in this paper tries to help.

➢ *Kenya*

A case study of Kenya is interesting in this regard. It boasts of a rich history of technology adoption and innovations, especially mobile money which has become an integrated component of economy, and has laid down a clear vision about itself in some policy statements. According to the Digital Economy Blueprint and the taskforce report on emerging digital technologies, 2019, AI and distributed ledgers were selected as focus areas (Ministry of ICT, Innovation and Youth Affairs, 2019). Kenya Digital Master Plan, 2022-2032 makes digital infrastructure, services, skills and innovation the key pillar of development in the country (Ministry of ICT, Innovation and Youth Affairs, 2022), and National AI Strategy, 2025-2030 was developed thereafter (Ministry of ICT and the Digital Economy, 2025).

The legal environment is important for the argument. The Data Protection Act of 2019 gives rights to individuals regarding their personal information. Section 35 states that the individual should not be subjected to any decision resulting from automated processing alone where such decision results in legal consequences or significantly impacts them, provided that some exceptions have been provided for (Republic of Kenya, 2019). What this implies is that any Kenyan bank, insurance company, employer or online lending organization is unable to consider the decision made by the agency as final in all situations. Some level of human intervention must occur for a decision that is significantly impacting.

Kenyan organizations also face practical conditions that shape agentic adoption. Data is abundant from the provision of credit through digital means and mobile finance services, with regulators having brought digital credit firms within their purview, making the consequences of automated decisions evident. However, in addition to most of these firms being small-scale, few people possess skills in artificial intelligence engineering and governance, and connectivity is uneven within the country. There is also the presence of informal economic activities that may not find representation within training data. With public institutions and county governments engaging in experiments in automation while operating on shoestring budgets and universities doing the same, it becomes clear that Kenyan organizations should not expect an existing agent to operate correctly in their environment. They must be careful about sharing decisions and supervising the agent.

➢ *What Makes AI Agentic*

The term "agentic" is often used in a vague manner, and therefore it's important to be explicit here. The agentic AI system in this article refers to one which, given a goal, is able to plan an action sequence, select and utilize objects from the environment, evaluate feedback and adapt behavior in a limited way without the involvement of a human user. It has been established in technical papers that language models have the ability to alternate reasoning and acting to achieve tasks which require multiple stages and outside information (Yao et al., 2023).

From the perspective of organization, what stands out is delegation. According to Baird & Maruping (2021), agentic information systems are different from the early forms of these tools by the fact that it is possible to delegate tasks to them, and in certain instances, the information systems delegate tasks to people. Conjoined agency is another related concept

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

by Murray et al. (2021). In conjoined agency, humans and technology share agency in acting on the world. Moreover, the manner in which this sharing happens influences how organizations do their activities. Taken together, these concepts suggest that the interaction between a manager and agentic systems is not tool-based but rather task-based.

There are varying levels of autonomy. In the case where an application generates the email and then awaits confirmation, the autonomy is limited. However, when the application goes through the complaint emails, reviews the customer accounts and refunds a certain amount before updating the case record, there is higher autonomy involved. In my definition of autonomy, it involves a sliding scale of autonomy from suggestions, actions with approval, actions with review after action and fully autonomous actions within mandate. This sliding scale is one dimension of the framework.

- *Decision-Making in Organizations when AI is Involved*

Shrestha et al. (2019) offer a useful framework for understanding the relationship between AI and decision-making processes. The authors classify decision-making structures into three types, namely full delegation to AI, hybrid structures where both AI and humans work in succession, and aggregation, where AI and human inputs are aggregated together. The classification is based on some characteristics of decision-making process such as the specification of decision-making space, alternative options and time constraints. As the authors argue, there is no universal structure that suits all types of decisions.

Another economic perspective offered by Agrawal et al. (2019) provides a complementing perspective. They suggest that AI reduces the costs of prediction and thus increases the importance of judgment, which is the selection and evaluation of outcomes. If true, agentic systems that use predictions would still be dependent on human or organizational judgment as to what goal to achieve and how to balance various factors. An agent could efficiently pursue a goal set for it, while some other person should set this goal, and if he or she failed to do that, efficient damage could be done. Brynjolfsson (2022) offers another caution as to how the development may proceed. If AI development focuses only on imitation and substitution of human activities instead of extending their capabilities, organizations risk losing their capabilities. This is called the Turing trap.

The augmentation question is not simple, however. Raisch and Krakowski (2021) show that automation and augmentation are intertwined and that organizations often face tension between them. If a firm automates a task, it may weaken the human skills needed to supervise it. If it only augments, it may miss efficiency gains. Dellermann et al. (2019) propose hybrid intelligence as a way to combine the complementary strengths of people and machines, so that the joint system achieves better results than either alone. These ideas support a design approach in which collaboration is arranged on purpose, not left to emerge.

- *Trust, Reliance and the Limits of Oversight*

The effectiveness of cooperation between people and agents is largely determined by human ability to establish trust and to follow the AI recommendations. Glikson and Woolley (2020) conducted a review of the experimental literature on human trust in AI and revealed that it has both cognitive and emotional elements and that it is determined by the tangibility, transparency and reliability of the artificial intelligence. According to Logg et al. (2019), people tend to prefer algorithmic recommendations to recommendations offered by other people, which the authors term algorithm appreciation. It matters since over-trust is one of the dangers of agent capability and fluency.

Findings from organizations also support this perspective. In their research concerning the use of AI systems in making diagnoses by medical professionals, Lebovitz et al. (2022) discovered that people had to put in effort in order to make a decision about interacting with the AI system whose basis of its output was not clear to them. Fügener et al. (2022) further revealed that combining AI with humans may result in poor performance if people are having trouble in assessing when they should depend on the AI, and the capacity of delegating tasks to the AI would help achieve better results. From the organization's point of view, Kellogg et al. (2020) demonstrate how the algorithms can be employed to supervise, assess and sanction employees.

Taken together, this research suggests that human oversight is a skill and a structure, not a checkbox. Green (2022) makes a similar point from a policy perspective. For agentic AI, where actions happen at speed and across several steps, the challenge is greater. A supervisor may only see the final outcome and not the chain of intermediate actions. Selbst et al. (2019) remind us that fairness and similar properties belong to the whole sociotechnical system and cannot be judged by looking at the model alone, which is a reason to design the human role and the organizational process together with the agent.

- *Gaps this Paper Addresses*

Three key gaps come to mind. Firstly, much of the literature on decision-making sees AI primarily as an advice-giver, so it does not address the issue of how decisions might change with the use of AI. Secondly, there is little literature on human oversight that considers different types of decisions with regards to which form of oversight is needed, rather than just whether a human should be involved at all. Thirdly, few studies exist on how organizational context in poorer countries may affect these decisions. Berente et al. (2021) call for research that looks at AI as a unique phenomenon and takes into account its autonomy, learning capacity, and inscrutability, and Mikalef and Gupta (2021) find that organizational capabilities in terms of AI, such as skills and culture, impact performance. The proposed framework in Findings attempts to bring these threads together.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

III. METHODOLOGY

➢ *Research Problem And Objectives*

The problem addressed is that organizations are starting to deploy AI systems that act, not only advise, while existing decision-making research largely assumes that the human takes the final action. The study has three objectives: (1) to clarify how agentic AI changes the nature of organizational decisions; (2) to explain what determines an appropriate division of labour between people and agents for a given decision; and (3) to identify the organizational conditions that make that division work, with attention to Kenyan conditions.

➢ *Research Design*

The study is conceptual and theory-building. It does not test hypotheses with primary data. It uses a purposive narrative review of the literature as its evidence base, followed by an iterative synthesis in which concepts from separate research streams were compared, combined and refined into a single framework. This approach suits an emerging phenomenon where empirical evidence is still thin and where the main need is for clear constructs and testable propositions.

➢ *Sources and Selection*

References were selected based on their relevance to four main topics – decision-making in organizations involving artificial intelligence, delegation of work to and collaboration with artificial agents, governance and regulation of AI technology, and digital transformation in the context of Africa and Kenya. References included peer-reviewed scholarly journals and conference papers, as well as policy and legislation documents published by the African Union, the European Union, and the Government of Kenya, together with some industry reports that could be used solely to describe the adoption of AI technologies. All references are dated between 2019 and 2026. The selection is not comprehensive but purposive, which means that the review does not aim at covering all possible relevant studies. Searches were conducted through Google Scholar within 2019-2026, combining keywords agentic AI, AI agents, decision-making, human-AI collaboration, delegation, information systems and digital transformation, along with Africa and Kenya. The main searches returned about 18,600 results for agentic AI and decision-making in organizations, about 17,700 for human–AI collaboration and delegation, about 323 for agentic information systems and delegation, and about 1,620 for AI and digital transformation in Africa and Kenya. Because these counts are too large to read in full, sources were selected purposively by relevance from the top results, supplemented by citation tracking and by official policy documents, so the review is selective and not exhaustive.

➢ *Framework Development*

The framework was developed in three steps. First, recurring features of decisions that the literature links to the suitability of AI involvement were identified and reduced to three: stakes, ambiguity and reversibility. Second, the degree of autonomy granted to an agent was separated from the type of decision, and four collaboration modes were derived from their combinations. Third, organizational conditions that the literature associates with effective AI use were grouped into four enablers. Propositions were then stated for each relationship the framework implies. Kenyan legal and policy documents were used to test whether the framework could accommodate local constraints, in particular the rules on automated decision-making.

➢ *Limits of the Approach*

Because the evidence is a purposive review and not a systematic one, other readers might select different sources and weight them differently. The propositions are therefore offered as claims to be examined, and the Kenyan material rests on documents and published research, not on fieldwork.

IV. FINDINGS

The main result of the analysis is a conceptual framework with four elements, which is presented below together with the propositions that follow from it.

➢ *Overview*

The framework rests on one central idea. The right way to share a decision between a person and an agentic system depends on the characteristics of the decision and on the autonomy the agent is given, and how well the chosen arrangement works depends on organizational enablers. It has four elements: decision characteristics, agent autonomy, collaboration modes and organizational enablers. Outcomes feed back into the first two through learning. I describe each in turn.

➢ *Decision Characteristics*

Three characteristics of a decision determine how much human involvement it needs. The first is stakes: the size of the potential harm or benefit to people, to the organization and to third parties, including legal effects. A decision to reorder office supplies has low stakes. A decision to deny a loan, dismiss an employee or release a patient has high stakes. The second is ambiguity: the degree to which goals are contested, information is incomplete or the right answer depends on values and context. Here I follow the economic point that judgment becomes more valuable as prediction becomes cheaper (Agrawal et al., 2019). The third is reversibility: whether an error can be corrected cheaply and quickly. A mistaken price update can be changed in minutes. A mistaken disclosure of personal data cannot be undone.

These three characteristics extend the features that Shrestha et al. (2019) used, and they fit agentic systems especially well because agents act in the world, so the consequences of error are real and sometimes lasting. In Kenya, the Data Protection Act adds a legal element to stakes, since decisions with significant effects on individuals attract duties relating to automated processing (Republic of Kenya, 2019).

➢ *Agent Autonomy*

The second element is the autonomy that the organization grants. Following the continuum described in the review of literature, I distinguish four levels. At the first level the agent suggests and a human acts. At the second the agent acts after human approval. At the third it acts independently,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

and a human reviews afterwards. At the fourth it acts independently within a defined mandate, with review by exception. Autonomy is a design choice made by the organization, not a fixed property of the technology. The same underlying system can be configured at different levels for different tasks.

➢ *Four Collaboration Modes*

A combination of decision attributes and autonomy leads to four ways of collaborating, as shown in Table 2 below. These are not fixed categories but regions in a continuum, and a particular process can encompass several of these.

The delegate mode is one where the agent decides and operates according to the mandate, which is human-defined and reviewed. It works best for decisions that have low stakes, low ambiguity, and high reversibility – things like allocating tickets, scheduling, or aligning routine data. The benefit here is efficiency and scale, and the only major risk is silent drift when there's no review.

Supervision in the supervise mode occurs where there is a human agent but humans watch and can interfere either by prior approval or review post-action. This mode is ideal for decisions that have medium stakes or medium ambiguity, and whose errors can be corrected but will become prohibitively expensive if compounded. This mode is exemplified in giving out modest refund amounts, preliminary examination of applications or drafting of regulatory submissions. The problem in this model is rubber-stamping because overseers end up endorsing all actions because it would not be practical to assess each action (Green, 2022; Fügener et al., 2022).

In the co-decide mode, the human and agent decide on a particular decision together. The agent will do the gathering of evidence, generation of choices, scenario testing among other things, while the human evaluates the values, context and consequences and decides. This is best used in decisions characterized by high ambiguity, including market entry, restructure or client negotiations. The danger in this mode is anchoring to the framing of the agent, and as such the organization should ensure the agent presents alternative choices.

In the reserve mode, the decision stays with humans. The agent may supply background information at most, and it does not recommend or act. This applies to decisions with very high stakes, low reversibility and heavy ethical content, for instance dismissals, disciplinary action or decisions affecting a person's legal status. Reserve is also the right default where the law requires meaningful human decision-making, which in Kenya includes certain decisions touched by section 35 of the Data Protection Act (Republic of Kenya, 2019).

Table 2 Collaboration Modes for Agentic AI in Organizational Decision-Making

Mode	Agent autonomy	Human role	Typical decision profile	Kenyan illustration
Delegate	Acts independently within a mandate	Sets mandate; audits samples and logs	Low stakes, low ambiguity, easily reversed	Routing customer queries at a telecom or utility help desk
Supervise	Acts with prior approval or rapid review	Monitors, approves, intervenes	Moderate stakes or ambiguity; errors correctable	First-pass screening of loan applications, with staff review of flagged cases
Co-decide	Generates options and evidence	Weighs values and context; decides	High ambiguity; judgment-intensive	County planning choices informed by agent-generated scenarios
Reserve	Provides background only, or none	Decides and bears accountability	Very high stakes, hard to reverse, ethically loaded	Dismissal or disciplinary decisions; decisions with legal effect on a person

➢ *Organizational Enablers*

For a mode to work as expected, there are four enablers that are similar to those used in determining the capability of organizational AI (Mikalef & Gupta, 2021) and managerial difficulties as identified by Berente et al. (2021).

Data and Infrastructure involves factors such as data quality and representativeness, dependability of system infrastructure as well as connections for the agent. However, in many cases in Africa, data deficiencies and infrastructure disparities may contribute to increased risks of poor performance of the agent for certain individuals or regions (Gwagwa et al., 2020). Governance involves mandate, escalation mechanisms, logging, audit trail and incident response. Additionally, it includes compliance with the data protection and sector regulations. In the absence of logging of the intermediate activities of the agent, any form of supervision becomes cosmetic. The human capabilities involve the ability of the individual to question, test, and override the agent's activities. Without the ability to question the agent, supervision does not occur. Culture covers norms about trust, responsibility and learning from errors. A culture that blames individuals for agent failures discourages reporting, while one that treats failures as information supports improvement.

➢ *Feedback and Learning*

In the end, the theory assumes that the collaboration process is dynamic. Results, close calls, and overrides provide evidence that can be used to fine-tune both the autonomy of the agents and decision making modes. This issue is rather straightforward. There will never be any organization which will always stick to the same decision making modes without changing anything. Autonomy can only be earned by agents through their performance and should be taken away from them if needed.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

➢ *Locating the Framework in Context*

The framework is intended to travel across settings, but the enablers will differ. In a large firm in a high-resource setting, the binding constraint may be governance or culture. In a Kenyan small or medium enterprise, it may be data and skills. In a public body, it may be legal accountability. The framework therefore does not prescribe one answer. It offers a structure for diagnosing which mode fits which decision and which enabler needs attention first.

➢ *Propositions*

The framework generates propositions that can be examined in future empirical work. I state six.

- Proposition 1. Organizations that shift towards co-decide or reserve approaches will face fewer serious decision failures as stakes and irreversibility escalate than organizations that keep agent autonomy high.
- Proposition 2. Ambiguity being increased will reduce the advantage of agent autonomy and increase the advantage of co-decide approaches, where humans will add judgment to context and values.
- Proposition 3. In supervise mode, the quality of oversight depends less on the presence of a human reviewer than on the reviewer's capability, authority and access to the agent's intermediate actions. Where these are weak, supervise mode will behave like delegate mode with an added appearance of control.
- Proposition 4. Organizational enablers influence the relationship between collaboration mode and decision quality; the more enablers are weakened especially regarding data representativeness and human capability, the smaller is the scope of decisions that can be entrusted.
- Proposition 5. Organizations that adjust autonomy level of their agents regularly based on historical outcomes and interventions outperform organizations that only determine the level of autonomy at the deployment of the solution.
- Proposition 6. In jurisdictions with legal requirements on automated decision-making, such as Kenya's section 35, organizations that design collaboration modes to meet those requirements from the start will face lower compliance costs and lower reputational risk than those that retrofit human involvement after deployment.

V. DISCUSSION

➢ *Theoretical Implications*

The framework offers contributions to the theory in three respects. First, it alters the debate regarding the nature of the issue – from whether AI must make a decision to whether the decision must be made in partnership, thereby going beyond the automation vs. augmentation distinction (Raisch & Krakowski, 2021). Second, it provides means for linking the kind of the decision and the kind of the human role, which addresses criticism towards the human-in-the-loop approach (Green, 2022). Third, it takes into account the reversibility factor, which is particularly important when dealing with agentic systems due to actions in the physical world.

It also relates to the notion of joint agency (Murray et al., 2021) and delegation theories of agentic information systems (Baird & Maruping, 2021), where the latter discuss types of joint agency, but the former provides a framework for making the choice between the two. However, it brings to light another challenge for theory development in the future. Since autonomy is bestowed by organizations and stripped by organizations, the accountability is left to people and organizations, not to the agent. Theories of organizational responsibility need to take into account systems that act.

➢ *Implications for Managers*

For managers, the lesson is to start from the decision, not the tool. First, understand the decisions that are in a particular process, assess their level of stakes, ambiguity, and reversibility, and assign a mode. Start with decisions that have low stakes, are reversible, in either delegate or supervise mode, collect data and scale up only as performance warrants. Always require logs for intermediate steps of the agent, otherwise supervision is not possible. Build up the capacity of those who will be supervising, and ensure that they have the authority to say 'no'. Lastly, beware of the Turing Trap (Brynjolfsson, 2022): if all deployments of agents are based on justifying that they will replace staff, then there may be a problem in eroding the necessary judgment required for co-decide and reserve decisions.

The warning sign from industry is leading in the same direction. If most agentic projects fail because of lack of clarity in their value or poor risk management (Gartner, 2025), the decision-first approach may allow avoiding the mistake of beginning with technology and trying to find its application afterwards.

➢ *Implications for Kenya and Similar Contexts*

For Kenyan organizations, three implications stand out. First, compliance and design should go together. Under Section 35 of the Data Protection Act, for those decisions that have great impact on individuals, such as credit, employment and insurance decisions, full independence of agent action cannot be considered as safe by default (Republic of Kenya, 2019). Organizations need to design for explanation and recourse right from the start. Secondly, data bias represents a practical risk. An agent that was trained on data that are biased against informal workers, rural residents or certain local languages will not be able to make good decisions on their behalf, hence the need to test agents on local data and keeping humans in the loop in cases where there is insufficient evidence (Gwagwa et al., 2020; Birhane, 2020). Lastly, capability development is a prerequisite and not an add-on. The AI strategy and the Digital Master Plan clearly indicate skills as being important (Ministry of ICT, Innovation and Youth Affairs, 2022; Ministry of ICT and the Digital Economy, 2025).

There is also an opportunity. Kenya's experience with mobile-based services shows that organizations can leapfrog legacy systems. Agentic AI could play a similar role in areas such as customer service, agricultural advisory and public service delivery, if organizations pair it with sensible governance. The African Union's call for AI centred on

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

African priorities (African Union, 2024) is best served when local organizations make their own decisions about how much to delegate, and on what terms, instead of accepting default settings designed elsewhere.

➢ *Implications for Policy*

In the case of regulation and policy making, the framework implies that any rule demanding a person in the loop should clarify what type of human input is required for what decision. An excessively vague requirement may lead to rubber-stamping. In this regard, the guidelines can refer to the concept of mode-to-stakes and reversibility, asking companies to justify their mode assignments and capabilities of reviewers. This approach will be in line with the risk-based reasoning behind the regulation adopted by the European Union (European Union, 2024).

VI. CONCLUSION

➢ *Summary*

Agentic AI changes the position of technology in organizational decision-making. It moves from advising to acting, and that move reopens questions about authority, responsibility and judgment that digital transformation research has not fully answered. This paper has proposed a framework in which the right form of collaboration depends on the stakes, ambiguity and reversibility of a decision and on the autonomy granted to the agent, and in which the success of any arrangement depends on data and infrastructure, governance, human capability and culture.

This thesis makes a modest claim that is, however, quite useful. The design question organizations face is not one of choosing either human decision-making or decision-making by machines; rather, the design question organizations face is one of how to divide decisions, and they are much more likely to find good solutions if they have a clear classification of decisions, gradual autonomy, accountability to humans, and the capability to monitor effectively. This design question has a legal and developmental angle as well for emerging countries such as Kenya since both the law and development strategy limit full automation of decisions. Empirical testing and refinement of the model proposed in this paper are the tasks ahead.

➢ *Implications*

This paper provides a method of selecting among types of shared agency, a “decision first” approach for managers, and a basis for identifying the type of human involvement that the regulator wants for a particular decision. The initial points of entry for Kenyan organizations are regulatory compliance within the scope of section 35 of the Data Protection Act, local data agent testing, and development of supervisory skills. This is discussed further in the Discussion section.

➢ *Limitations of the Study*

There are certain constraints that one must consider while reading this paper. This paper is conceptual in nature and the hypotheses presented herein have not been empirically tested. The literature review is dynamic in nature and there is no peer reviewed literature on the topic that is most relevant, technical and practitioner oriented from 2025 and 2026, and hence the use of literature is limited to this paper. The theoretical framework presented herein is oversimplified in nature. In practice, there is an interplay between the variables of stakes, ambiguity and reversibility which can even be contested within an organization and the four modes must be considered as regions of a spectrum. The study on Kenya is based on literature and policies only.

➢ *Future Scope*

Several lines of research follow. Case studies and field experiments could examine how organizations in Kenya and elsewhere assign decisions to modes and what happens to decision quality, speed and fairness. Survey research could test whether the enablers moderate outcomes as Proposition 4 suggests. Experimental work could study whether reviewers with access to an agent’s intermediate steps detect errors better than reviewers who see only final outputs, and how this interacts with algorithm appreciation (Logg et al., 2019). Comparative studies across African countries could explore how differences in regulation and infrastructure change the appropriate modes. Research is also needed on accountability, in particular how organizations and courts allocate responsibility when an agent acting within a mandate causes harm.

- Funding: This research received no external funding.
- Conflict of Interest: I declares no conflict of interest.
- Ethical Statement: This is a conceptual study based on published literature and policy documents. It involved no human participants and no primary data collection, so ethical approval was not required.
- Data Availability: No new datasets were generated or analysed. All sources are listed in the references.

REFERENCES

[1]. Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous intelligence for complex goals—A comprehensive survey. *IEEE Access*, *13*, 18912–18936.

[2]. African Union. (2020). *The digital transformation strategy for Africa (2020–2030)*. African Union Commission.

[3]. African Union. (2024). *Continental artificial intelligence strategy: Harnessing AI for Africa’s development and prosperity*. African Union.

[4]. Agrawal, A., Gans, J. S., & Goldfarb, A. (2019). Exploring the impact of artificial intelligence: Prediction versus judgment. *Information Economics and Policy*, *47*, 1–6.

[5]. Baird, A., & Maruping, L. M. (2021). The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. *MIS Quarterly*, *45*(1), 315–341.

[6]. Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, *45*(3), 1433–1450.

[7]. Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*, *17*(2), 389–409.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26oct568

[8]. Brynjolfsson, E. (2022). The Turing trap: The promise and peril of human-like artificial intelligence. *Daedalus, 151*(2), 272–287.
[9]. Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering, 61*(5), 637–643.
[10]. Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data: Evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.
[11]. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, 101994.
[12]. Eke, D. O., Wakunuma, K., & Akintoye, S. (Eds.). (2023). *Responsible AI in Africa: Challenges and opportunities*. Palgrave Macmillan.
[13]. European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*.
[14]. Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive challenges in human–artificial intelligence collaboration: Investigating the path toward productive delegation. *Information Systems Research, 33*(2), 678–696.
[15]. Gartner. (2025, June 25). *Gartner predicts over 40% of agentic AI projects will be canceled by end of 2027* [Press release]. Gartner, Inc.
[16]. Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660.
[17]. Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review, 45*, 105681.
[18]. Gwagwa, A., Kraemer-Mbula, E., Rizk, N., Rutenberg, I., & De Beer, J. (2020). Artificial intelligence (AI) deployments in Africa: Benefits, challenges and policy dimensions. *The African Journal of Information and Communication, 26*, 1–28.
[19]. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.
[20]. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366–410.
[21]. Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science, 33*(1), 126–148.
[22]. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes, 151*, 90–103.
[23]. Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., ... Clark, J. (2025). *The AI Index 2025 annual report*. Stanford Institute for Human-Centered AI.
[24]. McKinsey & Company. (2025). *The state of AI in 2025: Agents, innovation, and transformation*. QuantumBlack, AI by McKinsey.
[25]. Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management, 58*(3), 103434.
[26]. Ministry of ICT, Innovation and Youth Affairs. (2019). *Emerging digital technologies for Kenya: Exploration and analysis* [Report of the Distributed Ledgers Technology and Artificial Intelligence Taskforce]. Government of Kenya.
[27]. Ministry of ICT, Innovation and Youth Affairs. (2022). *Kenya digital master plan 2022–2032*. Government of Kenya.
[28]. Ministry of ICT and the Digital Economy. (2025). *Kenya artificial intelligence strategy 2025–2030*. Government of Kenya.
[29]. Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review, 46*(3), 552–571.
[30]. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210.
[31]. Republic of Kenya. (2019). *The Data Protection Act, 2019* (No. 24 of 2019). National Council for Law Reporting.
[32]. Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). *AI agents vs. agentic AI: A conceptual taxonomy, applications and challenges* (arXiv:2505.10468). arXiv.
[33]. Selbst, A. D., boyd, d., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 59–68). ACM.
[34]. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66–83.
[35]. Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research, 122*, 889–901.
[36]. Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems, 28*(2), 118–144.
[37]. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the International Conference on Learning Representations (ICLR 2023)*.