Agentic AI and Human Agency in the Future World of Work: Redefining Autonomy and Responsibility

Nubi Achebo
PhD, Nigerian University of Technology and Management Apapa, Lagos Nigeria

***Abstract*—Significantly, the development of agentic artificial intelligence systems, characterized by goal-directed, adaptive, and semi-autonomous action, has profoundly restructured the architecture of work and the foundational concepts of autonomy, agency, and responsibility. In contrast to previous automation systems, which were designed to assist and augment human decision-making, the development of agentic artificial intelligence systems has introduced the concepts of distributed and collective agency, human intentionality, and machine execution. This study examines the impact of agentic artificial intelligence systems on human autonomy and agency in the context of changing work arrangements. Through the use of theories and perspectives from the disciplines of philosophy, organizational studies, cognitive sciences, labor economics, and artificial intelligence governance, the study proposes a conceptual framework for the distinction of human autonomy, agency, and responsibility in the context of hybrid human-artificial intelligence systems.**

Through a comprehensive analysis of various governance models, task allocation mechanisms, skill transformations, and sectoral studies of the healthcare, manufacturing, and knowledge work sectors, the study finds evidence that agentic artificial intelligence systems are not only augmenting and replacing human labor, but are, in fact, transforming human agency itself. In addition, the study proposes various strategic scenarios of preservation, erosion, and drift to predict the future of human agency in the context of changing AI autonomy. In conclusion, the future of work would be determined not by the presence of artificial intelligence systems, but by the development of autonomy, agency, and responsibility in the context of collaborative human-artificial intelligence systems.

***Index Terms*—Agentic AI, Artificial Intelligence, Workforce, Human Agency, Human–AI Collaboration, Human Agency, Work Autonomy, Algorithmic Governance, Distributed, Responsibility, Trustworthy Artificial Intelligence, Organizational Design, Digital Transformation of Work, AI Ethics and Regulation, Future of Work.**

I. INTRODUCTION

Artificial intelligence tools are widespread in the contemporary work environment, and their use requires human involvement in decision-making to achieve organizational objectives. Therefore, there is a need to consider the relationship between human and agentic artificial intelligence in the work environment.

Research Question:
How and in what ways does the emergence of agentic artificial intelligence (AI) systems influence and inform the future of human autonomy and agency in the work environment?

There is an orthogonality between the concepts of machine agency and human agency, and between agency and autonomy (Suchman and Weber, 2016). This is especially true in the context of the Agentic-AI Foundation, which argues that agentic artificial intelligence systems can independently develop and perform goal-directed behavior. Agentic-AI systems describe a unique type of machine agency that transcends the conventional concept of digital automation and provides an important area of research on autonomy and agency in the context of agentic work systems. From this perspective, autonomy and agency are pressing issues, and thus one may consider the concept of responsibility. From the perspective of agentic artificial intelligence systems, there may be a need to develop novel forms of autonomy that were not conceptualized for the future work environment (Dahlin, 2023; Mitelut et al., 2023).

II. CONCEPTUAL FOUNDATIONS: AUTONOMY, AGENCY, AND RESPONSIBILITY

The word "autonomy" has Greek roots that mean "according to one's own law" (Dahlin, 2023). At its most basic, autonomy refers to "freedom from external control," "self-direction," "independence," and "the bearer or originator of actions." Autonomy includes decision-making capabilities, which are "the ability to choose in accordance with one's own desires."

Academic literature has clearly distinguished between the concepts of freedom and autonomy. Freedom, in the negative sense of the word, refers to the absence of interference, while autonomy refers to the ability to reflect, choose, and act on one's own motivation. An autonomous actor acts on reasons of his own choice and not on coercive drives, cultural scripts, or trends. In the work context, autonomy refers to the extent of decision-making authority vested in an employee, including the number of decisions, the restrictions imposed by others, and the approvals required (Langfred et al., 2016). Autonomy further refers to the ability to align work with personal values and goals.

III. AGENTIC AI: CAPABILITIES, LIMITATIONS, AND IMPLICATIONS FOR LABOR

Agentic AI refers to AI systems capable of performing autonomous actions as independent agents. This means these AI systems can carry out self-directed operations without explicit human intervention. The scope of operation of these AI systems varies, as some are confined to relatively closed task environments, where they can learn, adapt, and adjust their operations independently. This means these AI systems can perform operations independently without human intervention. Agentic AI systems, therefore, can impact human labour at both global and granular levels.

Agentic AI in the workplace thus represents a situation in which, on the one hand, human agency can be extended through collaborative processes, and, on the other, it can be supplanted if processes are fully automated and taken over by AI agents. The agency of AI agents is thus considerable, and it occurs within a scope narrower than that currently contemplated for agentic AI systems (Sapkota et al., 2025). Automation, as a rule, reduces the need for human labor, but a design focus on augmenting human agency can potentially increase productivity, income, and agency, thus offering an alternative route. The AI Agent Agency paradigm advocates a systematic mapping of labor agents deployed within agentic systems and a design that encompasses multiple levels of adjunctive AI, as suggested by Dahlin (2023) and Fanni et al. (2023).

3.1. Autonomy of AI Agents in Work Environments

There are concerns about the autonomy of AI agents, especially in collaborative decision-making, which brings about issues of responsibility, trust, and recognition of outcomes. In this regard, organizations are developing governance structures that include HR-related governance, transparency, and compliance (Grobler et al., 2014). It is argued that an equilibrium between governance and autonomy will help organizations benefit from AI agents. The future of AI agents suggests that, with further enhancements in design and knowledge, their autonomy will continue to impact organizations and boost business growth.

3.2. Human–AI Collaboration and Joint Agency

Joint agency refers to cases in which two or more agents cooperate to achieve a common goal. Human–AI collaboration refers to a joint agency in which humans and AI agents align on a common goal and work together to achieve it (Holter & El-Assady, 2024). Agentic systems differ in how they realize joint agency, ranging from replicating the entire workflow to replicating a single task, with or without human oversight. While AI systems that act as collaboration partners for humans are beneficial for joint agency, AI systems that act on their own are not beneficial for collaboration (Dahlin, 2023).

Shared goals and the ability to collaborate are the defining features of Human-AI Collaboration as a joint agency." In addition, the AI collaborator's ability to condition the inputs also plays a role in human-AI collaboration. Agent coordination occurs through the conditioning of inputs, which affects agents' behavior based on other agents' behaviour (Omicini, 2020). The AI collaborator plays a role in the design of the human-AI collaboration interface. The collaboration interface might be a module for agents to condition the collaboration, or for agents to collaborate through alternative means. Nevertheless, the decision on the

interface design has significant implications for the quality of the collaboration.

3.3. Accountability and Responsibility Across Agents
Agentic AI, the autonomous and intentional action of AI, includes the physical, spatial, and temporal aspects of activity in the work context. Agentic AI involves cognitive decision-making processes and the supervision, monitoring, and control of data by human and AI agents within the overall framework of joint agency. Agentic AI enhances human capacity, and although limited, human capacity is significant. For instance, in information work, the human-AI collaborative interface has limitations in meeting prescriptive, descriptive, and transformative needs, as suggested by Methnani et al. (2021). This has been supported by empirical evidence indicating the need for human-AI collaboration to be personalized, flexible, and interactive, and to include training for both human and AI agents, leading to the use of intelligent, agentic AI agents to augment human capacity.
The use of AI technology enables the agent to operate autonomously without external control, thereby enhancing its capacity, which would otherwise be limited by human assistance. This, therefore, leads to a decrease in organizational control, enabling the agent to co-author, enhance interactive and communicative materials, allocate time, divide tasks between agents and media, and make transformative changes, often beyond the operator's capacity (Venkatesh and Jayavardhan, 2025).
The absence of responsibility manifests as harm at the system level, with traceability but no causal links according to historical records (Jansson and Fasano, 2025). There may be the absence of interactive, progressive, and appropriate trials; the presence of informed agents performing wrongful acts; and the presence of system authority without sufficient dominance, such that individual responsibility may be uncertain and lead to systemic agents being held responsible. Theoretically, an agent may serve as a collective agent. Still, human agency is constrained by the need for regulation, which may result in the release of individual agency to agentic AI.
The digital shift in work, especially with the advent of AI, has revived the concept of work autonomy. Work autonomy is defined as the "extent of freedom provided to individuals to define both the subject matter and the manner of completing the task." Along with agentic AI and responsibility, work autonomy is another important concept in the study of human agency and agentic AI in the context of the changing nature of work. The intentional manifestation of agency, or the capacity to act and choose, in interaction with other agents and artefacts, depends on the establishment of autonomy at an appropriate level. It is thus important to discuss autonomy in AI-assisted work environments, as Dahlin (2023) proposes.
The digital age has seen increased discussion of autonomous work and its implications, particularly in the context of work mediated by AI systems that analyze data, provide recommendations, and perform tasks on behalf of the user across areas such as logistics, correspondence composition, predictive modelling, and software development, among others. To illuminate and guide discussions on the possibility of autonomous work and the changes and challenges that may accompany it, a working typology of autonomy in the digital age is proposed.

IV. TASK ALLOCATION, CONTROL, AND OVERSIGHT

To determine the proper allocation of task control and oversight for human actors and artificial intelligence in human tasks, several factors may be considered. First, the party that retains the final decision-making authority to execute the task should be determined. This includes the range of actions that the human actor and the artificial intelligence system can perform autonomously in relation to the task.

Second, the aspects of the entire task that are monitored or supervised by the human actor and the artificial intelligence system should be determined. This includes the whole task and its aspects, such as quality and impact.

Third, the party that retains final controlling authority over the information generated by the artificial intelligence system, as well as the assumptions and approaches it uses, should be determined. Higher levels of supervisory authority, within a broader framework, determine the nature of shared authority or influence (Niehaus et al., 2022).

4.1. Skill Development and Lifelong Learning for Humans and Machines

Hybrid human-AI systems are the result of the digitization of knowledge work. This requires the development of new human skills and the creation of new human tasks. Humans at different levels collaborate with artificial intelligence systems to develop the entire decision framework and monitor its execution. This delegation of human tasks to artificial intelligence systems affects human skills and the nature of the organization's work structure.

Human skills are still important despite the agentic nature of the systems. Large language models, such as ChatGPT, use generative conditional processing to handle large amounts of content when responding to user inquiries. User inquiries may come from people with varying levels of expertise, and as current AI models become more effective, the need for enhanced human skills, especially for advanced systems, is increasingly emphasized (Fanni et al., 2023).

4.2. Ethical and Legal Considerations

Agentic AI raises ethical concerns about value alignment, fairness, bias, discrimination, and more. The EU AI Act has established a set of high-level principles for AI. However, many questions remain about how the principles are interpreted and what the implications of agentic AI are for trustworthiness. This calls for the development of flexible reference points, standards for AI trustworthiness, and national implementations. The legislative framework is also lagging behind the rapid technological development of AI. This is seen in the development and application of AI, including liability for damages caused by AI, ownership of AI-generated work, treating AI systems as digital agents, and the personhood of AI (Lovell, 2023). The trustworthiness of AI agents' interests and their alignment with human interests are also unresolved. The development of AI systems, especially in the context of debates over human enhancement enabled by digital technologies, genetics, and pharmacology, also calls for many considerations of the framework. The debate is ongoing on enhancement, not on personal autonomy. The discourse on human rights is also influenced by the issue of AI agency (List, 2019).

V. MODELS OF GOVERNANCE FOR HUMAN–AI WORK SYSTEMS

The governance of heterogeneous human–AI work systems calls for the development of corresponding governance frameworks that guide and regulate agentic labour at the organizational and societal levels. This includes developing principles, guidelines, and institutions that foster the emergence of agentic organizations and support the development of agentic systems across sectors of the economy.

When developing standards, regulations, and compliance requirements, it is also necessary to propose associated governance standards, regulatory requirements, and compliance mechanisms that define and regulate agentic systems. This includes developing specifications for interoperability, safety, sustainability, and accountability. Moreover, it is essential to propose auditing systems that support organizational, national, or regional compliance with regulatory requirements while aligning with organizational goals, norms, and values.

In articulating principles of trustworthy AI, it is imperative to relate risk management with autonomy from a digital economy perspective. This includes developing requirements for reliability, explainability, and remediation while considering various levels of decision-making autonomy. A wide range of safety and security requirements should also be proposed to avert potential adverse consequences, especially in human–AI collaboration systems.

Regarding design principles and recommendations, it is important to initiate cross-sector discussions among organizations and job functions that use agentic systems. This entails analyzing changes in organizational hierarchies, workflows, and decision-making processes, including task allocation across multiple levels. Furthermore, it is important to address changing configurations, such as specialist, generalist, and hybrid profiles, in relation to task-distribution models. Moving forward with automation, particularly at advanced levels, standards, rules, and compliance mechanisms need to be established to support the human-AI division of labor. Human stewardship in relation to automation is an important factor in ensuring safety, well-being, accountability, and the promotion of autonomy in agentic technologies.

To achieve this, it is necessary to establish specifications that promote interoperability, safety, and accountability, taking into account the system and process dimensions, and utilize these standards in relation to technology development, compliance, and certification. In this context, there are significant challenges regarding the appropriate level of authority for agentic technologies and the appropriate method of human stewardship (Brendan Walker-Munro & Zena Assaad, 2022).

For example, in solid-state manufacturing, fully autonomous systems may be deployed where compliance with design and regulatory requirements can be conclusively established. Sustainability, health, and safety standards will dictate the appropriate materials and processes that may be used. The BASF compliance series may be an appropriate model for addressing these limitations in relation to production processes. In knowledge-based automation, human stewardship may be appropriate, particularly regarding confidentiality and ethical issues related to misinformation (Niehaus et al., 2022).

5.1. Trustworthy AI and Risk Management

The risks associated with AI systems include decision-making outcomes, security risks, and societal risks (Fanni et al., 2023; Venkatesh & Jayavardhan, 2025). The European Union's proposed AI Act has established a definition of trustworthy AI that draws on narrow and broad human contexts. The narrow definition refers to AI systems that comply with laws and requirements. Although this definition is consistent with the EU White Paper on AI and OECD principles on AI trustworthiness, it is not a comprehensive definition of the human AI interaction process, where humans operate cooperatively with an AI system that is agentic.

The second, more comprehensive, concept of trustworthy AI focuses on the objectives, values, and rights of humans. In the cooperative interaction between humans and AI systems, trustworthiness is based on improving the capabilities and knowledge of humans through comprehensive education and rationales for the actions of AI systems, a mission statement, absence of harmful effects, disputing the actions of AI systems, and monitoring. "Agency" is an important part of trustworthy AI. The continuous association of AI with humans, without a standard control protocol for the overarching premises and mission statement of AI, is highly risky.

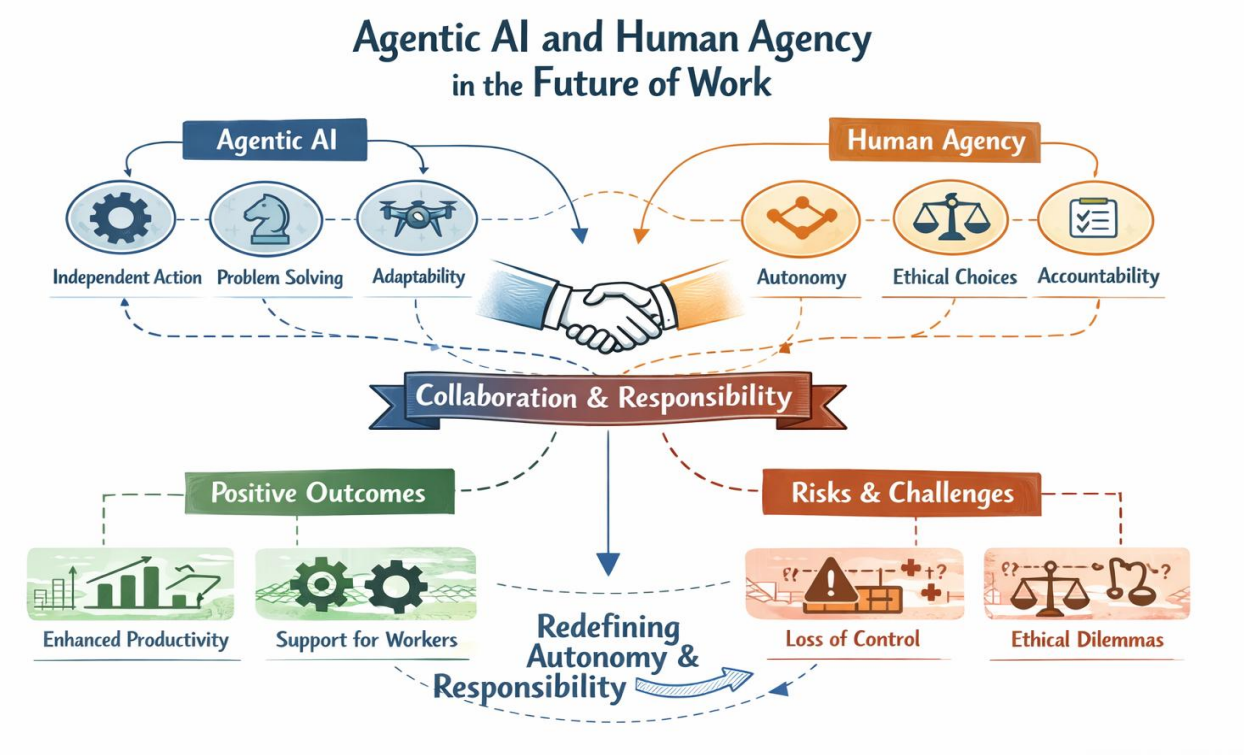

Diagram 1: Agentic AI and Human Agency

5.2. Organizational Design for Agentic Systems

The organizational design of an agentic system can significantly shape organizational hierarchies and work processes. In this context, the ability of an agentic system to concurrently interact with multiple agents can significantly impact organizational hierarchies, even with highly skilled personnel. Digital assistants connected to expert systems can significantly democratize access to knowledge, yet they can also raise trust issues.

As part of the human-AI co-creation strategy, the organizational structure may prompt an evaluation of potential changes to agents' and systems' creative specializations. In this regard, the collaborative, clinical, production, and knowledge domains might substantially impact the human-AI co-creation strategy. In this regard, the criteria for allocating decision-making authority could be applied across asset classes. In the broader policy context, the human-AI co-creation strategy might shape models of agency within the organizational structure, with significant implications for the balance of power and control between humans and systems, as Niehaus et al. (2022) highlight. The use of principle-challenging human-AI technologies might be subject to the precondition of questions of authority, cooperation, and trust, as highlighted by Methnani et al. (2021).

To determine the appropriate allocation of decision-making authority between the human agent and the artificial intelligence, several criteria may be used to evaluate the human-AI co-creation strategy. Firstly, the human agent might be responsible for the final decision-making authority for executing the tasks. In this regard, decision rights might be defined as the extent of autonomous action permitted for the human agent and the AI system in the execution of tasks. Secondly, the human agent might be responsible for monitoring various aspects of the tasks, such as their quality, progress, and impact. In this regard, the extent of the granted authority may be a significant variable in defining the roles of human agents and AI systems within the overall process, as highlighted by Methnani et al. (2021). Thirdly, the human agent might be responsible for controlling the information, assumptions, and decisions used by AI systems to execute tasks. In this regard, the supervisory role might be a significant variable in defining the human-AI co-creation strategy, as highlighted by Niehaus et al. (2022) and Murkhejee and Chang (2025).

VI. CASE STUDIES: SECTORS RECONFIGURING AUTONOMY AND RESPONSIBILITY

Healthcare organizations are increasingly using agentic systems to redistribute decision-making autonomy from clinicians to patients, while maintaining clinicians' responsibility. Patient-centric AI systems forecast patients' needs based on individual histories, thereby facilitating coordination of subsequent steps in diagnostic and treatment processes (Lundberg et al., 2021). These systems autonomously formulate propositions and requests for data, with the clinician responsible for the propositions.

Manufacturers are considering the potential to use agentic systems across the entire range of production processes to redistribute responsibility for quality in line with new technologies' capabilities, thereby facilitating the development of autonomous factories. These systems autonomously manage production processes, including the formulation of requests for human consideration and approval (Dahlin, 2023). Responsibility for the critical objective of quality rests with the operator, but this allows for prioritization of attention to more challenging aspects.

Knowledge-intensive organizations are using agentic systems to augment, rather than replace, human decision-making, thereby maintaining responsibility while achieving productivity gains. These systems are used to support decision-making throughout the decision life cycle, including curation, consideration, selection, creation, evaluation, and justification, autonomously and collaboratively with humans.

Table 1: 20 Global Cases of Agentic AI Deployments

Organization	Brief Description of Deployment
Amazon	Autonomous warehouse robots that navigate, pick items, route packages, and collaborate as multi-agent systems.
Google DeepMind	Agentic AI controlling data-center cooling systems, autonomously optimizing energy consumption.
Tesla	Autonomous driving agents making real-time navigation, lane-change, and hazard-avoidance decisions.
Siemens	Industrial AI agents optimizing factory lines, machine settings, and predictive maintenance schedules.
JPMorgan Chase	Agentic AI that detects fraud, assesses financial risk, and autonomously triggers protective transaction actions.
Airbus	AI agents predicting aircraft component failure and automating fleet health management.
Walmart	Autonomous shelf-scanning robots detecting out-of-stock items and updating inventory systems.
IBM Watson Health	Clinical decision-making agents assisting diagnosis, treatment suggestions, and patient triage.
Baidu Apollo Go	Autonomous ride-hailing vehicles navigating urban environments with agentic driving systems.
Alibaba Cainiao	Multi-agent logistics AI orchestrating parcel routing, sorting, and autonomous vehicle operations.
Microsoft Security Copilot	Agentic cybersecurity AI autonomously detecting threats and generating real-time mitigation strategies.
UPS	AI agents autonomously optimizing delivery routes to reduce time and fuel consumption.
Shell	Autonomous refinery monitoring agents detecting anomalies and initiating safety actions.
Unilever	Supply chain agents autonomously adjusting production planning based on real-time demand signals.
Netflix	Personalization agents autonomously optimizing user recommendations and content sequencing.
Pfizer	R&D AI agents autonomously screening compounds and accelerating drug discovery workflows.
Meta	Content moderation agents detecting harmful content and autonomously removing high-risk posts.
HSBC	Compliance agents autonomously scanning transactions and documents for AML violations.
Uber ATG	Autonomous driving agents for ride-sharing fleets executing sensing, mapping, and navigation tasks.
John Deere	Autonomous tractors and harvesters making real-time decisions on routing, spraying, and soil optimization.

6.1. Healthcare and Patient-Centric AI

AI has been integrated into the healthcare industry through strategic regulations that ensure patient autonomy and safety by promoting accountability and transparency through effective policies and architectures. The discussion focuses on the potential for collaboration between humans and AI and highlights a shift in focus from the human-patient interface to the human-AI-patient interface. The healthcare industry's interest in developing health

technologies and addressing the ethical issues associated with them has been acknowledged. Organizations have examined the potential of chatbots and other technologies to improve efficiency and safety without compromising autonomy.

6.2. Manufacturing and Autonomous Production Lines

Autonomous systems redefine the implementation of production processes, task allocation, quality management, and defect detection. Autonomous systems perform the roles previously assigned to human workers. This includes the acquisition of skills previously regarded as a human monopoly. Autonomous systems manage production output per hour in line with market demand. They ensure the sustainability of the supply chain and maintain the right inventory levels to prevent disruptions. The main aim of implementing robotic systems is to assist human workers in completing multistage assembly. This includes assessing the components' compatibility using the digital twin. The twin is responsible for procuring the necessary equipment and identifying potential hazards. However, the issue of quality management, whether in human-robot or autonomous systems, remains a major challenge for implementing human-AI cooperation (Sidorenko et al., 2023).

Agents capable of performing roles set by humans raise questions of safety, responsibility, and regulation. Deficits in energy supplies, global economic instability, international conflicts, and the spread of pandemics threaten the global value chain. Industries in the process of change, retraining their workforce, and acquiring new labor forces are more vulnerable to risk than ever. Autonomous systems redefine the scope of human agency. This is because they autonomously allocate the scope of responsibility for the assigned task beyond human capacity. This leads to the creation of a new boundary based on the scope of their operation (Dahlin, 2023).

.

6.3. Knowledge Work and AI-Augmented Decision Making

Information-intensive jobs that have traditionally relied on human expertise for creativity, judgment, and insight are increasingly leveraging AI systems for supporting decision-making and cognitive productivity. Large language models such as ChatGPT, Claude, and Gemini allow users to access natural language responses for complex questions and contextual information. AI systems are also useful for scrutinizing lengthy texts, summarizing them, and making inferences. AI systems are also useful for generating original texts across genres and styles, writing poems and lyrics, and creating outlines for various types of writing.

The involvement of AI systems in these activities blurs the lines of human and AI contributions for tasks that require creativity and insight. In these activities, human workers also contribute significantly by prompting AI systems, reviewing outputs, and critically analyzing them to synthesize information from diverse sources. In these activities, human workers are likely to experience a lack of accountability because of the involvement of AI systems. However, developing ethical guidelines for these activities is an important issue that has been left to future research (Dahlin, 2023).

VII. STAKEHOLDER IMPACTS: EMPLOYEES, EMPLOYERS, AND SOCIETY

Agentic AI poses new questions about the future of human labor, especially regarding human agency and attributing responsibility for AI-related actions. The theoretical and practical importance of these questions is reinforced by the widespread adoption of AI systems by the enterprise sector. The fundamental transformation in the nature and form of work across workplaces is already apparent in perception, knowledge, language, and creativity.

Agentic AI refers to an AI system capable of self-directed action, including goal-setting, planning, and execution under constraints. The AI system can affect its environment and other agents through information, physical objects, and materials. An organization selects an agentic AI system, and the power to enhance, alter, or modify the agent is under the organization's control. Digital autonomy and organizational agency are distinct concepts, with each agent responsible for its actions within its respective domain. The various configurations of an agent depend on how control and oversight, skill development, accountability, and autonomy relate to each other. To manage risks and ensure that opportunities associated with agentic AI systems in the workplace are maximized, a framework should be adopted concerning introducing the AI system,

monitoring its activities, and performing actions based on the initial rationales for its introduction (Nedzhvetskaya & Tan, 2021; Niehaus et al., 2022).

7.1. Labor Market Transitions and Inclusion

Changes in the labor market are caused by agentic systems and other enabling technologies, affecting inclusivity. Training systems may not be able to keep up with the changing workforce due to agentic systems and other technologies. Even if training is agile, it may not be sufficient to determine whether in-demand jobs exist or whether workers have the required skill set (Dahlin, 2023). The ability to be flexible and create more occupational opportunities may be an advantage. Pay gaps have also widened between the micro-workers exploited by agentic systems (Tubaro & Casilli, 2019). Dependent tasks and discriminatory pricing may cause workers to exit the labour market (Frank et al., 2023). The consequences of agentic systems need further exploration. Accessibility challenges affect workers' adoption of technology, and market inaccessibility impedes people with disabilities' adoption, causing equity issues. The open approach to thousands of workers with disabilities has the potential to accelerate the adoption of assistive agentic systems, thereby increasing autonomy, agency, and safety during job transitions.

7.2. Worker Agency, Empowerment, and Safety

Agentic AI, defined as AI systems that can act autonomously in changing environments, raises issues of worker agency and empowerment, especially when workers are physically removed from the site of activity. The emergence of agentic AI in one area of the labor market may undermine authority, autonomy, and agency in another (Fanni et al., 2023; Leonardi, 2025). Agentic AI may also raise issues of worker safety and well-being. The nature of the workplace environment varies by job type, and digital documentation of the environment sheds light on workers' experiences of encountering agentic AI. Analysis of the data has shown that agentic AI may increase the risks faced by workers who are subject to task replacement by other agents and may impede workers' engagement with safety risk assessments (Nedzhvetskaya & Tan, 2021). The undermining of worker safety and agency may lead to increased anxiety related to job loss (Dahlin, 2023).

VIII. AI FRAMEWORK: INTERDISCIPLINARY FRAMEWORKS

The combination of law, economics, cognitive science, and organizational studies has the potential to yield a more comprehensive understanding of the complex relationships among humans, AI, and agentic responsibility (Hughes et al., 2025). The regulation of human–AI interaction, including private law liabilities, rights, auditing, and traceability, is an additional aspect that requires consideration (Dahlin, 2023). The economic perspective must also consider, among other issues, the influence of agentic systems on the adoption of technology, work organization, and power distribution, as well as historical theories of technology, work, and the economic transition of sentient systems. Cognitive developmental theory, along with system dynamics, must also be used for conceptualizing AI systems, human–AI interaction, and AI systems as agents in work systems (Fanni et al., 2023). Organization studies must also be used to explore, among other issues, the complex relationships between power, hierarchy, agency, and freedom in human–AI interaction, as well as boundary issues for agentic systems and hybrid agency.

IX. STRATEGIC SCENARIOS FOR THE FUTURE OF WORK

Key factors and indicators must be determined and analyzed. The resilience strategies must also be assessed. The strategic scenarios must include technology transition paths based on government policies. The strategic scenarios must also include paths based on government research initiatives and simultaneous investments in education and social safety nets.

9.1. Optimistic, Pessimistic, and Aberrant Pathways

Different strategic scenarios should be defined in terms of the divergent evolutionary paths of the work context. These should include an optimistic path in which the development of agentic AIs supports human agency and welfare, a pessimistic path in which the development of AIs fails to support human agency and welfare due to the poor development of AIs, and an aberrant path in which human agency and welfare are undermined in the context of the development of AIs, along with the safety issues involved. The major

factors influencing the strategic scenarios should be defined in terms of their indicators. The strategies for building resilience in the strategic scenarios should be discussed. The strategic scenarios should be defined in terms of technology development paths driven by government policies, research initiatives, and investment in education and social safety nets. The defined strategic scenario paths are as follows:

1) Agency Preservation Scenario: This strategic scenario is an optimistic scenario in which human agency, governance, technology, and society utilize the enabling potential of AI to enhance human agency. Incremental explorations and impact accelerations are discussed in the context of the resilience literature (Niehaus et al., 2022). The indicators of the strategic scenario are the development of foundation models for human-AI joint agency (Mitelut et al., 2023) and the consideration of the societal impact of AI by stakeholders (Dahlin, 2023).

2) Erosion Scenario: This strategic scenario is expected to involve considerable risk and a lack of understanding of agentogenesis, value alignment, and policies influencing human-AI cooperation. The indicators of the strategic scenario are the reevaluation of human-AI collaborations with significant impact and the termination of AI collaborations by individuals who cannot protect and utilize them.

3) Drift Scenario: In the context of multi-agent systems, the concept of joint agency is defined as the shared circumstances in which agents are influenced beyond their capacity. However, the concept of joint agency may be poorly represented. Large language models may be capable of producing substantial outputs that influence circumstances beyond their capacity. However, the capacity for distributed joint agency may be limited.

9.2. Policy-Driven Scenarios and Interventions

Proactively addressing adverse pathways that may threaten an individual or society is essential. The necessary policy interventions should be considered with a view to leveraging the substantial opportunities that agentic AI offers to achieve positive outcomes. The aim is to foster positive pathways by developing innovative best practices that benefit society, while addressing adverse pathways that may impede development or undermine positive ones. Transition plans should be developed that incorporate all essential elements, including capacity-building interventions, effective risk-mitigation strategies, and changes to roles, relationships, and operations. The capacity to adapt to anticipated changes should also be built.

9.3. Roadmaps for Organizations

Organizations should develop strategies that enable the adoption and assimilation of agentic systems, aligned with regulatory and policy frameworks (Dahlin, 2023). A measured approach should be adopted that recognizes the capabilities and limitations of agents and the interplay between human and non-human agency. The potential paradoxical impact on organizations and sectors should also be considered, and actions and interventions should be analyzed in terms of their position on a spectrum between opportunity and vulnerability. Organizations should develop staged transition plans to assimilate agentic work systems, aligned with investments and complementary capabilities across organizational functions. The range of capabilities that influence vulnerability to and opportunities from agentic systems should also be considered, with a focus on targeted use-case selection (Niehaus et al., 2022).

X. CONCLUSION

Autonomy is no longer an individual prerogative in the execution of tasks. Instead, autonomy has become a complex, negotiated concept distributed across human actors, algorithmic systems, and organizational structures. Similarly, agency has emerged as a relational concept between human actors and AI systems, in which the actions of AI systems condition those of human actors, and vice versa. Moreover, the concept of responsibility has transcended the event-based notion of individual decisions to become a systemic concept embedded within socio-technical architectures.

Sectoral analyses of healthcare, manufacturing, and knowledge work domains suggest that agentic AI systems may both enhance and undermine human agency, depending on the nature of the system's governance and oversight. The positive "Agency Preservation" scenario illustrates AI systems' capacity to enhance human agency when appropriately governed and overseen. At the same time, the "Erosion" and "Drift" pathways suggest the undermining of human agency due to inadequate

oversight and skill development. The primary claim of this article is that the future of the world of work will not be defined by the technological capabilities of agentic AI systems, but by normative and institutional decisions regarding control, transparency, redress, and capability development.

Therefore, human agency in the future of the world of work, dominated by agentic AI systems, requires appropriate oversight, skill development, control structures, and governance to ensure the trustworthiness of AI systems. Redefining autonomy and responsibility is not an abstract philosophical exercise but an imperative that requires organizational and institutional interventions. The future of the world of work, dominated by agentic AI systems, will not be defined by the autonomous capabilities of machines but by the capacity of human systems to manage autonomy responsibly.

REFERENCES

[1] Dahlin, E. (2023). And say the AI responded? Dancing around 'autonomy' in AI/human encounters. ncbi.nlm.nih.gov.

[2] Fanni, R., Eveline Steinkogler, V., Zampedri, G., & Pierson, J. (2023). Enhancing human agency through redress in Artificial Intelligence Systems. ncbi.nlm.nih.gov.

[3] Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., ... & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. Proceedings of the National Academy of Sciences, 116(14), 6531-6539.

[4] Grobler, A., Bezuidenhout, M. L., & Hyra, A. (2014). Governance and HR: the development of a framework for South African organisations. Journal of Contemporary Management, 11(1), 164-184.

[5] Holter, S., & El-Assady, M. (2024). Deconstructing Human-AI Collaboration: Agency, Interaction, and Adaptation. Computer Graphics Forum, 43(3). https://doi.org/10.1111/cgf.15107

[6] Hughes, L., Dwivedi, Y. K., Malik, T., Shawosh, M., Albashrawi, M. A., Jeon, I., ... & Walton, P. (2025). AI agents and agentic systems: A multi-expert analysis. Journal of Computer Information Systems, 65(4), 489-517.

[7] Jønsson, T. F., & Fasano, M. C. (2025). Theorizing subjective responsibility at work: an agentic approach. Frontiers in Psychology, 16, 1548931.

[8] Langfred, C. W., & Rockmann, K. W. (2016). The push and pull of autonomy: The tension between individual autonomy and organizational control in knowledge work. Group & organization management, 41(5), 629-657.

[9] Leonardi, P. M. (2025). Homo agenticus in the age of agentic AI: Agency loops, power displacement, and the circulation of responsibility. Information and Organization, 35(3), 100582. https://doi.org/10.1016/j.infoandorg.2025.100582

[10] List, C. (2021). Group Agency and Artificial Intelligence. Philosophy & Technology, 34(4), 1213–1242. https://doi.org/10.1007/s13347-021-00454-7

[11] Lovell, J. (2023). Legal Aspects of Artificial Intelligence Personhood: Exploring the Possibility of Granting Legal Personhood to Advanced Ai Systems and the Implications for Liability, Rights and Responsibilities. Rights and Responsibilities.(May 10, 2023).

[12] Lundberg, J., Arvola, M., & Palmerius, K. L. (2021). Human autonomy in future drone traffic: Joint human–AI control in temporal cognitive work. Frontiers in Artificial Intelligence, 4, 704082.

[13] Methnani, L., Aler Tubella, A., Dignum, V., & Theodorou, A. (2021). Let Me Take Over: Variable Autonomy for Meaningful Human Control. Frontiers in Artificial Intelligence, 4. https://doi.org/10.3389/frai.2021.737072

[14] Mitelut, C., Smith, B., & Vamplew, P. (2023). Intent-aligned AI systems deplete human agency: the need for agency foundations research in AI safety. arXiv preprint arXiv:2305.19223.

[15] Mukherjee, A., & Chang, H. H. (2025). Agentic AI: Autonomy, Accountability, and the Algorithmic Society. arXiv preprint arXiv:2502.00289.

[16] Nedzhvetskaya, N., & Tan, J. S. (2024). The role of workers in AI ethics and governance.

[17] Niehaus, S., Hartwig, M., Rosen, P. H., & Wischniewski, S. (2022). An Occupational Safety and Health Perspective on Human in Control and

AI. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.868382

[18] Omicini, A. (2020). Towards a notion of agent coordination context. In Process coordination and ubiquitous computing (pp. 187-200). CRC Press.

[19] Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges. Information Fusion, 103599.

[20] Sidorenko, A., Motsch, W., van Bekkum, M., Nikolakis, N., Alexopoulos, K., & Wagner, A. (2023). The MAS4AI framework for human-centered agile and smart manufacturing. Frontiers in Artificial Intelligence, 6, 1241522.

[21] Suchman, L., & Weber, J. (2016). Human-machine autonomies. Autonomous weapons systems: Law, ethics, policy, 75-102.

[22] Tubaro, P., & Casilli, A. A. (2019). Micro-work, artificial intelligence and the automotive industry. Journal of Industrial and Business Economics, 46(3), 333-345.

[23] Venkatesh, A. N., & Jayavardhan, G. V. (2025). Future of Work: Managing Ethical Challenges of Agentic AI and Super Intelligence in Organizations. Advances in Consumer Research, 2, 5277-5284.

[24] Walker-Munro, B., & Assaad, Z. (2023). The Guilty (Silicon) Mind: Blameworthiness and Liability in Human-Machine Teaming. Cambridge L. Rev., 8, 1.