Deep Science Publishing
https://doi.org/10.70593/978-81-988918-1-5_3

DeepScience
Open Access Books

Chapter 3: Understanding agentic artificial intelligence: Autonomous digital agents and their impact on workflows and decisions

3.1. Introduction

What if an AI chatbot could operate a task entirely on its own in software and/or hardware? What if, when you asked your favorite AI chatbot to schedule a meeting with someone, it created your calendar appointment, contacted the person, negotiated the time slot, and sent reminders, handling any situation that arose, entirely without your further input? These so-called "autonomous" or "agentic" AI models are rapidly emerging. In this chapter, we explain how autonomous AI systems already exist now and provide an overview of what they can do. We also discuss the implications this new AI paradigm could have for work and decision-making (Bryson et al., 2020; Acharya et al., 2025; Hughes et al., 2025).

We frame agentic AIs as general-purpose computing systems that are able to perform an open-ended set of tasks in various software and hardware environments on behalf of or rather than just in support of human users. Central to our conceptualization of autonomous agents are three properties: the ability to act (and finalize tasks) on behalf of users rather than just in support of users; the ability to execute a wide variety of different tasks that are not limited to query-response or dialog functions; and the ability to work and/or interact in the physical world as well as in virtual environments. We define Agentic AI as those technologies and applications that are based on Large Language Models and their variants that exhibit the properties of such autonomous agents. Commercial "Agentic AI" systems and apps utilize also other AI-centric approaches including multimodal or evolutionary learning and/or hybrid models that integrate LLMs with neural-symbolic, cognitive architecture, or common sense

reasoning components. Initial projects in the relatively newly emerging field of Agentic AI focus either on using LLMs to augment existing automation technology or on autonomous process optimization and decision-making. But fully enabling Agentic AI would involve a significant redesign of current processes to fit into an autonomous model that assigns agencies to both humans and agents, a design that allows agents to perform their task autonomously with assistance from human users only when necessary and that allows human users to monitor agent operations and interventions at low cost (Stone et al., 2016; Russell & Norvig, 2020; Sapkota et al., 2025).

3.2. Defining Autonomous Digital Agents

The term agentic AI refers to a certain class of artificial intelligence that has the capacity to independently fulfill goals. Historically, Western Philosophy has identified rational autonomous agents as unique to humans.

Fig 1 : Autonomous Digital Agents

Yet numerous Eastern Philosophies espoused the existence of autonomous agents beyond humans. Regardless of their originator, humanity has endowed various artifacts, both mundane and sophisticated, with autonomously agentic properties such as the volcano, the wind, the tree, and the clock. More recently, the invention of digital computer technology has empowered humans to create digital artifacts imbued with varying degrees of agency, leading to the emergence of robotics and agentic AI. Just as robots are physical agents, agentic AI or digital agents, as semantic digital agents are better known, are the newest digital agents capable of acting on behalf of users, organizations, and society. Digital agents autonomously perform tasks in partial or full fulfillment of a user's goal. Digital agents manifest a goal-directed temporal cycle of perceiving the world, selecting appropriate actions, and taking actions in the world. The temporal cycle in which digital agents operate varies from seconds or minutes in the case of micro-tasking to days or months in the case of macro-recurring tasking.

Digital agents comprise: bots, which are software programs that communicate across a network and perform rudimentary natural language or other non-verbal tasks on behalf of a user; AI-enabled bots that perform elementary decision tasks on behalf of users and organizations by conducting keyword search and analysis; and service bots, which interact with back-end service systems to perform user-initiated or trigger-initiated non-compound micro-tasks. While bots, AI-enabled bots, and service bots are primitive digital agents, they exist in the digital world and society and perform user-centric services. Systems combining digital agents and AI-enabled physical robots can undertake and provide more advanced and responsible services to society and industry.

3.2.1. Characteristics of Autonomous Agents

An autonomous agent is a system that operates independently to achieve a goal. Although autonomy is a fundamental feature of these systems, the degree of autonomy can differ significantly. For example, an agent may have only limited decision-making autonomy, such as customizing website content for each visitor based on their web browsing history. Or it may operate with complete autonomy, such as an asteroid exploration probe traveling millions of miles in space and relying on its own sensors and knowledge to assess and direct the quality of the local scientific investigations.

Autonomous agents range from simple automated systems programmed to execute the same limited set of functions to sophisticated systems driven by advanced artificial intelligence. They differ from other intelligent computer systems designed to function as tools to be used by people to accomplish a task together. For instance, the various types of online services are generally designed to help users find the best answers to specified queries or provide intelligent support to decision processes led by people. But

autonomous agents take on full responsibility for doing things without any direct intervention by human operators.

Some autonomous agents work on behalf of a business or other institutional owner, such as an automated stock trading program programmed to operate independently in accordance with a set of pre-established rules. Other agents are developed as general-purpose service providers, performing specific tasks for a fee in a variety of categories such as online travel planning, entertainment reservation, retail shopping, and stock trading. In contrast to a simple automated system, these agents have an interactive, dialog-driven interface that enables them to converse with users, clarifying details about specified goals. They also use advanced artificial intelligence for natural language processing and similarity-reasoning functions to enhance communication capabilities.

3.2.2. Types of Autonomous Digital Agents

Some digital agents are so small or so closely coupled to the user that it is impractical to consider them separate from the user's mind. In these cases, it is best to use software tools that amplify the user's efficiency, memory, or cognitive ability. A second type of agent occupies the same cognitive space as the user but attends to some separate aspect of the task. These digital agents function as office assistants, working on some part of the task that may be time-consuming or difficult for the user. For example, a digital agent might undertake the task of filtering and classifying incoming email messages, flagging those that require the user's immediate attention, discarding spam, and learning the user's email classification and response patterns over time. A third type of agent is one that could be removed from the situation without changing the user's available cognitive capacity. These agents operate completely independently and can perform a function on their own. The more capable of intelligent operation an independent agent is, the more profound its impact on the user will be, credit in crises, watch-dog functions in surveillance systems, or monitoring tasks or other lengthy processes.

Various groups of researchers have developed different, often contradictory taxonomies of agents based on differing agendas and research paradigms. Agents may be classified based on one or more dimensions, such as the environment of the agent, the nature of the communication and cooperation between agent and environment, institutional aspects of the agent, the structure or internal architecture of agent systems, the organization of the agents, the learning and learning algorithms employed, and many other dimensions to build several different organizing schemes. These organizing schemes are procedural in nature, describing the agent's protocols, as well as structural, describing the agent's model. Moreover, agents can be classified as intelligent or non-intelligent agents, with the former using some kind of artificial intelligence techniques to accomplish an intent. Computer programs employed by human users to seek

information or carry out repetitive routines in a semi-automated or automated fashion are also sometimes considered agents. Such programs can send messages to one another but do not act with a user's intent.

3.3. Historical Context of AI Development

Intellectuals, futurists, and entrepreneurs have long posited the possibility of creating a machine with human-like capabilities: the creation of an entity that would deeply understand, reason, plan, and learn from experience. These clever thinkers and creators did not just contemplate this substantial intellectual feat, but many of them dedicated their careers to developing the mathematical, engineering, and neuroscience principles underlying agency and intelligence, and dedicated their time and ingenuity to the development of artifacts that would, over time, exhibit human-like levels of competence. The fruits of their lifelong labor are increasingly becoming ubiquitous.

The field of Artificial Intelligence (AI) is vast, overlapping with several disciplines. Cognitive Science, Control Theory, Computer Vision, Decision Theory, Internet of Things, Game Theory, Game Development, Independent Components Analysis, Linguistics, Machine Learning, Neural Networks, Non-Linear Dynamics, Parsing, Pattern Recognition, Philosophy of Mind, Psychology, Reasoning Theory, Robotics, Simulation, Speech Recognition, Theoretical Computer Science, etc. There exist specialized journals and conferences in many of these disciplines that are not specifically labeled as AI. Very simply put, researchers in these disciplines not only seek to understand how intelligent beings function, but they also seek to develop artifacts capable of performing intelligent tasks. AI includes both the study of intelligence, the development of intelligent artifacts, as well as the dynamics of such interactions. Historically, the term AI has been used more to describe behaving intelligently than understanding or building intelligent systems. AI has existed since the dawn of computing: in its earliest incarnations as expert systems, it was a practical necessity in specific domains.

3.3.1. Evolution of AI Technologies

With ancestral technologies such as machine learning, knowledge representation, and reasoning, Natural Language Processing, and Computer Vision, Artificial Intelligence were purposed and initially envisioned to become Digital Agents to augment and become involved in human-centric decision-making processes. Such idealization over the fertile ground of Information Technology development and the shortcomings of human-centric workflows to a species-wide extent has elevated users' expectations of AI capabilities to heights thought as being close, if not equal to, human intelligence. Amid these

circumstances, advances in computational power and the formulation of relevant inductive biases has lead to the prominence of Statistical Learning and Generative Learning Advances in both learning formalisms and algorithms to train on domain-expert designed dedicated Deep Neural Networks the rich modalities of digital data with stride.

The enormous success of Statistical Learning and Generative Learning Advances technologies fueled the revamping of the Data Architecture of existing non-AI workflows with Stochastic Network Architectures to incorporate their subsidization of human-centric decision-making processes, heralding an era of modern AI. The Robotic Process Automation Non-AI-to-AI Paradigm is the core theme upon which Conversational Agents, Intelligent Digital Agents, and Intelligent Digital Twins have been implemented, novel classes of software agents the main socio-technological contact with modern AI. The flexibility and customization capacity to address problem domains in a variety of user-oriented application landscapes proclaimed the acceptance of Non-AI-to-AI Paradigm-purposed technologies as the four pillars of modern AI. These agents are today deployed en-masse and considered synonymous with AI.

3.3.2. Milestones in Autonomous Agent Development

Agentic AI has evolved through decades of research into autonomous computational entities. Some of these early developments, characterized by expert systems and robotics, were not considered agents because of their limited connectivity and scope. Nevertheless, agentic AI relies on classic AI building blocks: perception through perception channels and actuators, reasoning through domain models, and acting in the world to attain goals.

The first agent-like systems appeared in the 1960s, with LittleDog and Eliza. The 1970s saw the first systems that gave the illusion of agency: interactive narratives such as Colossal Cave Adventure and planning neuro-symbolic models. In the 1980s, a new class of interactive narratives called interactive movies implemented some visual and structural characteristics essential to narratives: real-time graphics, interactivity, multiple storylines, and an overall plot. The late 1990s introduced agents as bots working online, and virtual autonomous agents working in the game industry. These agents had connections between modules, which is a critical property of modern agentic AI. In the early 2000s, personal assistant agents and chatbots could become the first conversational agents, bots engaging in text conversations with users. Smart agents implemented an open, scalable agent architecture for the Smart 2020 metaverse.

Now AI development has come full circle. Generative AI is the culmination of decades of successive advances in the key research areas of AI: algorithmic thinking,

computational creativity, machine learning-based economy, open AI technology, and model-based modeling as fourth kind of science. Generative AI also consolidates decades of market establishment where whole sectors of industries studied the application of AI-based systems. Agentic AI is next.

Fig 2 : Major Enofin Autonomous Agents

3.4. Technological Foundations of Agentic AI

Machine learning is the technological foundation of Agentic AI. The advancement of neural networks and, in particular, of deep neural networks have made it possible to achieve unprecedented results in the area of predictive learning. Deep neural networks have enabled the successful deployment of predictive models across a number of key domains such as: vision; speech recognition; language understanding; and semi-structured and structured prediction. These models are making their way into production and being used at scale, as evidenced by the rapid commercial success of multiple large tech companies in the arenas of vision and speech recognition, as well as the efforts of the broader AI community in the areas of language understanding and semi-structured and structured prediction.

Natural Language Processing is of particular importance to the area of Agentic AI because it constitutes a critical interface between human and machine agents. As most humans use natural language as the means of expressing their intentions to one another, it is only natural to use natural language as the primary interface between humans and Agentic AI systems. Recent advancements in the area of natural language processing owe a great deal to deep learning research, and — much like other areas of predictive learning — have enabled the creation of a rich set of predictive "building blocks" that allows for a great degree of transportability of solutions across tasks and domains. Specifically, deep neural networks pretrained on data from a diverse array of tasks and domains are being increasingly utilized for performance on task- and domain-specific natural language processing tasks by transferring knowledge from the pretrained model through a relatively small amount of task- and domain-specific training data. As with other areas of AI, natural language processing is in a state of rapid development, and a number of exciting challenges and opportunities lie ahead.

3.4.1. Machine Learning and AI Algorithms

Agentic AI corresponds to a juxtaposition of various trends in the field of mainstream AI and ML research. It is mostly thought of as a collection of self-optimizing agents developed by humans that are able to build knowledge and continuously autonomously train and deploy themselves on top of knowledge bases with the goal to solve problems by themselves. In order to efficiently act, these agents need to develop a representation of the complex real-world problem that they are dealing with, adapt to changing contexts and without being explicitly programmed, solve long-term problems. This conjunction requires the integration of various issues and sub-problems, such that a holistic integrative research agenda is created. It is not sufficient to specialize on one narrow issue, but rather develop sufficient solutions for various interdependent problems.

Almost all Agentic AI works so far have relied heavily on manually or semi-manually designed AI algorithms. These include, for example, reinforcement learning for policy specification, evolution strategies for exploring the policy space, deep neural nets for representing the policy, symbolic AI for knowledge representation, and reinforcement learning to optimize complex behavior compositions, or model-free hierarchical reinforcement learning for behavior optimization over complex behavior trees that can be learned using behavioral cloning or imitation learning. Other work has relied on the agent using external interfaces, for example, to access and request world data via sensors and provider interfaces, and use symbolic logic to learn and reason about existing knowledge.

3.4.2. Natural Language Processing

Natural Language Processing (NLP) is the set of processes and algorithms that allows computers to read, process, comprehend, and produce outputs in human language. With recent advances in large language models, the application of NLP in agentic AI has flourished. Still, there is a set of requirements NLP systems must address to bridge the gap toward higher agentic levels. Language, as a medium of human-agent communication, is a vital element of agentic collaboration.

Theoretical developments from the philosophy of language, natural language communication models, epistemology, and social epistemology can help advance NLP technology and algorithms. These developments can help deepen natural language comprehension and advance common-ground building, which is still rudimentary. Enhancements to NLP technology will allow human-centric workflow steps between agents and humans to be more natural, more efficient, and higher quality by creating additional opportunities for semantic grounding, joint attention, aligning beliefs, opinions, desires, and intentions, and collaborative referencing. Bridging and semantic recognition gaps in communication via other modalities can help enhance and expedite lower-bandwidth joint communication by resourcing additional communicative modalities collaboratively. The agentic NLP enhancement will help enable higher-level agent competencies that require advanced natural language abilities.

Agentic levels are the levels at which language is involved in the intelligent exercise of agency. At a first, lower level, NLP serves to help explain or criticize agentic work by humans and agents alike through feedback, bi-directional explicability, and causal argumentation. At a second, intermediate level, high-competent agents make vultures above work by deferentially algorithmicizing human tasks, such as requests for assistance, ethical, quality, and goal-checks, and quality assurance, and collaborating, with other agents, to automate the performance of those tasks.

3.4.3. Robotics and Physical Agents

Autonomous robots and physical agents represent a native tier of Autonomous Digital Agents (ADA) development. Modern Agentic AI began in the digital space with purely virtual ADAs handling workflows, data, and processes within other design patterns. Discrete actions by ADAs impacted the real world via digital inputs into physical machinery and onto real-world businesses, laws, and reputations. Meanwhile, since the late-twentieth century emergence of than-at-all-like Agentic pre-ADAs, early interest grew in robotics as a method for moving the output of the cloud-based ADAs into the material world. Recent advances in both ADAs and robotics have begun to reintegrate these nodes. Still other forms of robotics also leverage the same core technologies as

digital ADAs to change how we locally interact with the material universe without creating an obvious movement of labor from existing humans.

Physical tasks collide with training ML and RL algorithms more challenging than for digital ADAs. Specialized sensor-tech connects these robots to physical data from their local area. Safety requires GAIs undergoing intense social testing to manage the risk of causing fatal accidents. Yet here again, individual robots and connected networks of them can create not-DSA GDML services, with custom jobs like cleaning buildings or warehouses and lawn-mowing, delivery-robot functions. Research also continues into remote-agental robotics systems. Under AI's influence, these have begun tackling functions in dangerous far-flung locations at industrial scale, whether searching for landmines, archaeological surveys, or addressing oil spills.

Development and delivery of these local, agentically-intelligent ADAs benefit from both cloud-computing and dispersed robotics with their autonomy. The robots contribute by attempting non-highly-skilled tasks as tools or with partial autonomy. Human workers are involved to cope with ambiguity and solution, safety, and social-response problems. These early exploratory phases of evolution and integration toward Democratic Digitization take form for robotics and physical agents from without and within.

3.5. Applications of Autonomous Digital Agents

The modern era is full of countless automated systems and processes. Nowadays, people are subjected to a significant influence from algorithmic decision-making or algorithmic neutrality that is far more impactful than at any other time in human history. From search engine optimization to social media post generation to flying drones to supply chain management optimization, digital agents are omnipresent and their impact on lives and workflows will only continue to increase. The research domain of digital agents is inherently multidisciplinary, drawing influences from fields such as natural language processing, multi-agent systems, information retrieval, computer security, psychology and behavioural modeling, game theory and cognitive modeling. The goal behind the creation of agents is to develop systems that can perform specific tasks automatically for the user and that are capable of efficiency, autonomy, and reactivity to overcome the limitations of current automated systems and processes to help boost human productivity.

Traditionally, these agents have been often deployed for specific domains e.g. spell checking and grammar checking and sometimes, only for specific use-cases such as sentiment analysis. However, the advent of conversational agents and more specifically, task-oriented dialog systems and dialogue state tracking challenges has opened the door for new kind of intelligent conversational agents capable of tackling the massive world

of user queries from numerous diverse users in an open and continuous manner. Digital agents in terms of data analysis, strategy decision-making, and also alongside engineering, have been around for strategic decision-making for a long time. Their impact on workflow-related and decision-related automation tasks is about to significantly increase due to the advances being presently made in large pretrained transformer-based architectures, which are capable of unsupervised transfer learning on massive corpora and are gaining immense popularity and traction in the agent community.

3.5.1. Customer Service Automation

Autonomous Digital Agents (ADAs) are autonomous applications interacting directly with users to perform a range of online and mobile tasks. More specifically, ADAs are in direct customer interaction roles. In customer service settings, ADAs increasingly act in conjunction with human workers. In these settings, physical and cognitive task efforts are completed in combinations between consumers, Digital Agents (DAs), and human fulfillment agents. It is important at the outset to clarify some differences between ADAs and related digital applications. ADAs differ from chatbots and voicebots in that customer-centric Data Agents are open ended and conversational goal-seeking applications used for enhanced human problem solving. By contrast, chatbots are prompt-response matching Q&A systems based on heuristics and flows, while automated voicebots use speech recognition and output to simulate human interaction. DAs differ from ADAs in that data-centric assistants focus on pattern-matching and heuristic solutions rather than dynamic conversation and goal-seeking processes. Shortcomings of traditional customer service tools, such as chatbots, mobile apps, and human workers, provide a rationale for investment in ADAs.

First, customer service response management systems demand significant company resources. Customer complaints, support requests, and inquiries consume time and employee resources as firms receive increasing volumes of incoming queries. Information about the customer — including their browsing history, previous interactions, and order details — is often critical to resolving issues. Personalized services are often necessary, because response times and service outcomes are cornerstones to successful customer interactions. Yet many customer inquiries are repetitive and straightforward, such as service activation questions, routine questions about delivery status, website navigation requests, and password resets — all tasks that could be automated. For these reasons, organizations routinely engage customer service agents, bots, and assistants to answer low-complexity queries.

3.5.2. Data Analysis and Decision Making

Despite the rising popularity of AI assistants, most users employ these technologies to perform discrete, untrackable operations that do not create data that can be synthesized for insights. However, there are a variety of interactive, cloud-based, agent-focused programs that can be used as data analysis tools, creating a behavioral data trail that can deliver business intelligence on employee workflows and cognitive processes. Rather than being an automated tool that publishes a report, these digital workplace assistants can be invoked to create processes to query restored stored data on any topic, and publish human-like reports on that data or inform a user's decision-making flow regarding that data. Such reports are preferable to reports generated by business intelligence visualization tools, because the vast majority of stakeholders, consumers, and business decision-makers do not have deep experiential knowledge of statistical analysis or have advanced data visualization and summarization literacy. These limitations are barriers to actively engaging in the analysis of data about products or markets stored in business intelligence tools.

Generative LLMs can be employed as interactive decision aids by frontline employees as they analyze the predictive data for specific decisions embedded in dashboards using company dashboards developed by business intelligence automation tools. Such Agentic Decision Aid systems allow the AI assistant to engage with the user like an assistant, and have them explain specific variables being tracked on dashboards. Unlike traditional AI tools, with pre-written scripts and limited natural language capabilities, these systems allow users to provide immediate feedback about what areas are helpful or unhelpful. This allows them to communicate with business users in a natural, unscripted, and effective manner, guiding the decision-making process supported by embedded predictive insights.

3.5.3. Workflow Optimization

Enterprises are already starting to adopt agentic AI across various domains such as IT help desk functions; knowledge management; end-user computing, deployment, and monitoring; customer service; automated data analysis; and reports generation. As companies make heavier investments in IT and enterprise decisions — whether logistics-driven or demand-driven — be it infrastructure or cloud-based or SaaS products, agentic AI can help start to optimize workflows. Higher degrees of automation within workflows not only bring down the operational costs, improve overall service quality, enhance enterprise productivity, and create a near risk-free environment, but enable organizations to make these investments more cost-effective, customer-centric, agile, and able to adapt fast to changing market conditions requiring different products and services for demand generation. Agentic AI excels in first-order optimizations for enterprises —

understanding APIs connecting enterprise SaaS products to one another, monitoring and creating connections between them through the APIs to optimize workflows, calling functions, acting on the functions internally and externally, planning and scheduling actions, and even building rudimentary agents for multi-step functions. Although the current use of agentic AI in external automation of workflows is very limited because of its still-expensive usage, future advances in LLMs and other developments will surely bring down the costs leading to scaling such automation. However, organizations will be able to use agentic AI more freely in terms of internal process optimization and account management, where the activity space to explore is bounded using internal employment policies.

3.6. Impact on Workflows

One of the most profound effects of Agentic AI is its ability to carry out entire business processes without the need for management. Homogenous digital processors designed by programmers are already used in almost all transparent workflows, such as pick-up and delivery work by truck drivers from one location to another. This work is built around the negotiation of a price, which is driven by supplier cost structures, lack of demand-response mechanisms among suppliers, and the need for just-in-time deliveries at the lowest cost. Robotaxis will transport passengers, driverless trucks will transport cargo, and remote-controlled drones will transport small packages. They will be dispersed and will create demand-response mechanisms through pricing just the same as trucks and taxis do now. There are many more business processes that are boring but essential to the functioning of modern business. AI has for a long time been invading this space, and Agentic AI will continue this trend by pushing into areas where human error could result in large losses of human life or assets, laws regulating the financial services industry, medical decision-making, or routine air travel.

In a sense, Agentic AI creates the ultimate performance by reducing transactions, simple and normal decisions, and ordinary information management to the lowest cost by making the process invisible and incidentally invisible in large, long-standing businesses. It not only capitalizes on existing knowledge systems but invests in the creation and updating of knowledge weapons. These weapons will shape new knowledge systems that elevate the great majority of businesses to their optimal cost, service, and profit levels. This effect on performance means that workers will only need to intervene in businesses at the periphery, where human imagination and energy are the ultimate rare process. This enhancement of human work production by AI may be fruitfully compared to the effect of the advance in mechanization in the 19th century.

3.6.1. Streamlining Processes

Organizations engage digital agents to carry out specific tasks within governance frameworks established with respect to those tasks. Enabling agentic capabilities securely delegates ownership of completion of specific parts of a process to the agentic entity, which can be its creator, external, or internal to the organization. Processes are thus now constructed from independently executable and verifiable parts linked end-to-end to form a coherent whole. This transfers the burden of design of specific segments, and ownership of execution risk, onto the agent, which must remain within resource budgets for the allocated task. As the number of actors increases, human involvement in interactions necessarily grows. If we delegate and partition tasking, humans take on supervision. If we choose to combine tasks into multiparty interactions, humans take on facilitation. In either case the combination of actions through human inputs and agentic elements that design, execute, and verify raises challenges to establishing and maintaining reliable workflows. Humans are often bad at concentrating solely on task choices and do a passable job at deciding how best sequences should be accomplished when business as usual. It is at or near milestones, and when new problems are found, that we use our unique capabilities as facilitators to examine choices and modify workflows, cognizant of context, to cater for most current events and anticipated outcomes. This makes interactions multifaceted in both direction and timing using different media over time and provides scope both to be unsynchronized and to run in parallel. Agentic systems relieve facilitation pressure, permitting increased device-to-device processing. This will heighten productivity and lower reaction times ameliorating congestion, but human facilitators must still be present at key moments. Users will aid increase agentic efficiencies and integration errors. More efficient process flows will change when we execute workflows, how quickly we react when new events dictate modification of trajectories, and how we, and agents, communicate about these stimuli.

3.6.2. Enhancing Collaboration

Communication, cooperation, and collaboration are inherent properties of human life, and for a long time shared the unique attribute of being possible only among humans, at least as far as we understand. Collaboration — defined as the act of working together with one or more individual(s) to produce or create something — requires recognition, trust, exchanging information and knowledge, and building and agreeing on shared goals. More advanced forms of collaboration additionally require problem-solving, creativity, considering others' perspectives, and assessing and interpreting non-verbal cues. Computers have engaged in "collaboration" with humans as partners in the orchestration of tasks, such as during tandem cycling, but the computer has always been the operator not one of the collaborators. Agentic AI creates a new opportunity for digital

agents to enter into this domain. The crucial aspect here is the property of agency whereby a computer program presently regarded as no more than a glorified spreadsheet is transformed by an innovation within cognitive science into a digitally-created life, or more aptly into a digitally-created agent capable of communicating and collaborating with humans and the world we inhabit with them.

What would it look like for these digital agents, endowed with agency, to collaborate with us in the traditional sense of people working together? The truth is we don't fully know. A large portion of collaboration is driven by the specific domain and task at hand. But here are some initial thoughts: Many activities at the fringes of human action might see agents become valuable collaborators and partners. For example, actions that require working with schedules, such as booking vacations or scheduling appointments for contractors, are human actions with relatively high dollar values yet low money velocities. It is highly likely that some hybrid form of AI collaboration — agents performing preliminary browsing and recommending actions for signoff and action — would prove to be the best-defined design approach.

3.6.3. Reducing Human Error

Humans are not perfect. While those who undergo extensive and costly training usually perform jobs better than others, they are still prone to forgetting important steps or making mistakes for other reasons. Factors such as cognitive load, stress, boredom, exhaustion, and emotion all contribute to lapses in an agent's performance. These factors are especially relevant to work involving repeating tasks or rote memorization—often tasks that agents by design do for cheaper and faster than other people. Luckily for everyone else, this does not eliminate the risk of error in work processes for all workforces in all scenarios. It does, however, significantly decrease the chances of oversight. It might also increase confidence in unattended workflows. But keeping the human element convinces us who rely on agents that it is completely foolproof.

We have already alluded to the relationship between agentic AIs and the risk of human error during decision-making. An agent-enabled number crunching task might reveal some too-numerical data that needs a person to positively approve a general conclusion where the algorithm says to go ahead. Implementing agentic AI as a safeguard would mitigate that error-making element and act as a break for something needing cautious reflection rather than cold calculation at the helm. With certain jobs being more tedious than others, and those jobs requiring intensive repetition with focus and patience, vulnerabilities in various workman make them subject to fault at certain times of day. These factors make supporting agentic AIs suitable for many repetitive work and overtime scenarios such as manufacturing and security. Or applications like industrial sensors, process alerters, and network monitoring or debugging.

3.7. Decision-Making Processes

A major development area of applied AI can be found in corporation automation, in business orchestration systems that embrace the intellectual processes of those corporations and the enablement of intelligent agents by using cognitive capacities that were always associated with humans, but that can be automated for decision-making and action execution. Additionally, society has been adopting Decision Support Systems that assist their users on important issues, from flight on time scheduling decisions to credit possibilities for loan requests, being entrusted more and more as notice guides for sensitive decisions.

If the use of Decision Support Systems in day-to-day operations is a good example of possible action delegation to an idiotic system, the solution for specific constraints helps humans in understanding better what they are attempting to do, complicated decision-making processes are an inevitable consequence of the exponential increase of digitalized information and its underlying interconnections, opening space for agentic features and responsibilities in situations like strategic priority definition and selection of policies focused on long-run company perspectives, from negotiation strategies to collectors and refining plans, but also for policy comparison considering social responsibility, environment impact, risk, and contributions to the digital divide, paving the necessity of trust and explanation metrics for artificial decision-making as well as the combined use of AI models and humans to guarantee strategic vision within organization long-run decision processes.

The evolution of Decision Support Systems by introducing agent-based systems tackles the responsiveness and proactivity assistance for human decision-makers, but it has deep implications on strategic organization design. Incorporating intelligent systems on the top of the organization, defining responsible decision-making and the connected liability, will impact the definition of power and profit flow within an organization.

3.7.1. AI in Strategic Decision Making

Strategic decision making (SDM) has been defined as the process of making decisions involving the organization as a whole, and having destructive consequences that cannot be reversed. Decisions taken in this process shape the direction of the firm for the foreseeable future, and the extent of decisions taken is so vast, it is impossible to predict the impact in the long term, causing it to be described as a “black box”. However, in the AI literature, several have noted that SDM is not a unified, solitary act, but rather a process comprising several types of decisions. These varied decision types range from macro to micro aspects, and can be handled by varying levels of the organization. Each sub-decision serves a purpose, and together these purposeful sub-decisions help

organizations respond to internal and external stimuli. In the context of state-owned enterprises tasked with balancing political and economic goals, strategists have found that what is called "strategic decision making" at these firms is a collective process based on a series of tucked-away, make-do decisions that are often non-rational.

SE-Agents and organization-level AI can impact SDM directly or indirectly. Through their formal or informal power structure, organizations arrange information flows towards the people who are meant to act on them. This is done through politic and bureaucratic structures, so that people focus on different aspects of organizational actions. Chatbots – specialized SE-Agents – are expected to support micro-influencing decisions, ranging from answering queries to the IT department or book traveler arrangements among others. A new trend is generative AI, that has exploded in a projected market in a short time, with these tools augmenting the micro-decisions of employees throughout firms. These tools, when employed when employees are unsure of how to solve a problem, assure that the quality of the mini-decisions is higher than without this augmentation; that employees spend less time coming up with creative solutions; and minimize the number of micro-decisions that have to be revisited and redone later due to poor execution or changes.

3.7.2. Ethical Considerations in AI Decisions

The ethical implications of AI decisions are under intense scrutiny as autonomous systems are enabled to influence people's lives and increasingly tasked with performing sensitive decision-making activities, from prescribing force to people accused of crimes to recommending reliable news articles. The debate on trust in AI is fueled by two conjoined claims: proponents claim that the incapacity to understand the reasoning behind the AI's decision can lead to ever-increasing trust, while proponents of explainable AI urge that transparency and intelligibility of reasoning behind AI decisions is a necessary precondition for trust. In the context of AI in operations management, the two claims can be reconciled if AI algorithms affiliated with the human agent are developed with responsibility, fairness, and justice for impacted stakeholders as guiding normative principles.

In this section, we summarize some of the most important ethical principles considered in the current debate on how best to make AI responsible and trust-evoking decision makers. First, an agent – be it human or AI – must be able to justify its normative reasoning and help a critique understand how it reached a specific normative decision. Second, since responsibility is a form of consequentialism, the underlying utility function is crucial; it must be aligned with the moral values of the stakeholders impacted by its decisions. This can be difficult to achieve when users may sacrifice accuracy and reliability: for example, a selective AI striving to avoid accidents may decide to prohibit

all hairpin bends in snowy regions, or an AI checking for cancer might prefer to miss a cancerous blemish rather than labeling a healthy patch as suspicious. Also, even if the AI wants to do the right thing, a conflict of interest can arise when the gatekeepers want to optimize other, possibly conflicting, objectives. Fourth, in some sensitive domains individuals have a right to being treated fairly by employing the same criterion for similar cases. Finally, as biased algorithms may leave more harm than benefit when accessing vulnerable individuals or communities, ethical AI research participants and engineering regulators should take precaution.

3.8. Challenges and Limitations

Although Agentic AI has shown breakthroughs in task or role-based digital agent functionality, there still exist relevant limitations that could hamper their development and widespread adoption. Not all tasks are good candidates for Agentic AI. Humans and businesses still have their tasks that are not easily made agents of automation. This is due not only to cultural or organizational inertia but also to entrenched workflows requiring at least human-in-the-loop approaches. For larger-scale processes or projects where tasks have multiple interacting users, the complexities go beyond technical possibilities, such as switchovers of agency between different involved humans and Agentic AI. For other tasks requiring high reliability in predictive capability of outcomes, Agentic AI still utilize them with no guarantees from consumers and communities. Because this is the case, it is important that professionals consider the Agentic AI capability spectrum when considering for tasks these agentic services may augment.

Another challenge for the understanding of Agentic AI, as well as any automation technology, is the ethics of deploying autonomous decision-making systems. Algorithms are known to harbor bias that, if replicated into Agentic AI and deployed in critical tasks, may reflect and exacerbate systematic opportunity inequities. Non-trivialities of human-like probabilistic understanding is the possible infusion of human biases learned or re-informed by content and decision-making strategies across advanced models, and the difficulty in tracing and explaining their operationalization, especially in dynamic and scaled workflows. Compounding assessment difficulty is the black-box nature of some machine learning models used within an Agentic AI agent. Methods of differentiating algorithmic risk are still nascent in AI development teams, businesses employing inspection and audit, and third-party assessor practices. These challenges are further compounded in thin-market areas for assessment of ethical design and development of AI product. Finally, the possibility of Agentic AI service use for illicit purposes make these independent services attractive, yet exposed to sanctions or closures.

3.8.1. Technical Limitations of Current AI

There are several technical limitations of current AI that must be addressed before agentic AI systems can become commonplace in high-stakes decision-making contexts, including a lack of common sense reasoning, opaque models, a lack of integrated multimodal perception, and a lack of foresight and planning capabilities. Humans seamlessly integrate prior knowledge learned throughout their lives with current context to win chess matches, manipulate objects, identify sources of humor, and maintain long conversations. Current AI systems lack such common sense reasoning capabilities, which limits the kinds of tasks they can perform and the degree of confidence we can have in their outputs. This limitation will need to be addressed before they can be fully trusted in high-stakes applications.

Modern transformer-based language models are complicated. They may use the most sophisticated techniques, the biggest models, and the largest datasets, but at their core they are tuned statistical methods. Large language models are also opaque. While controlled prompting can induce desired behaviors, there remains an element of uncertainty. Their inner workings cannot easily be inspected. They cannot be simply queried to determine which prompt patterns will induce a desired behavior. It can be difficult to know how to interpret their outputs and there are many failure modes with no easily identifiable causes. These limitations remain challenging even after immense engineering and work to create these models. There can be a lack of transparency about how they come to a particular output.

AI systems make a concerted effort to avoid offensive outputs. However, it is important to keep in mind that these systems are not themselves ethical judgers. They have been trained on massive datasets that are known to include problematic content. Thus, it is impossible to guarantee that they are free of bias and harmful outputs. While such outputs may be rare, it is possible that they can find their way into high-stakes workflows. The only known solution is to ensure that human monitors are present in the decision-making loop.

3.8.2. Bias and Ethical Concerns

People are often unaware of the biases built into the training sets of current AI systems, leading to the promulgation of the stereotypes hidden there. When the stereotypes are negative to a group, it leads to the group continuing to be underrepresented and businesses opting to work with a model that is unknowledgeable about certain areas. A prominent example of this was when an image search system was unable to distinguish people with skin color darker than Caucasian. If an African American user searched for pictures of people with skin color darker than Caucasian, the pictures returned dominated

with images of spider-webbing. Accompanying this perception problem is the reality of defining and addressing the issues surrounding bias and ethical concerns. In general, there are two sets of problems. The first set of problems deals with the algorithms, whether directly or indirectly. The fact that AI systems used an algorithm to help find out that questions offered by wallpaper manufacturers reflect a user's housing color scheme is an example of direct consequence.

The second set of problems handles the fact that no matter what system one has deployed, the world continues to evolve underneath us. For example, it is known in Group Theory that one can have abnormal characteristics, but produce normal answers. However, there could be ethical issues where the opposite is observed. With AI systems currently deployed in commercial entities, it goes without saying that this should be corrected as much as possible. For example, the above situation would preclude systems designed to monitor and track social media. However, such systems have become increasingly common, for better or worse. The first set of ethical issues related to algorithms has been addressed numerous times in the past.

3.8.3. Regulatory and Compliance Issues

The application of Agentic AI in various fields brings about concerns surrounding applicable laws and regulations, primarily because current laws are not equipped to handle Agents as actors in decision-making processes. One of the first considerations is whether current companies using Agents in their workflow are in breach of fiduciary duties, international standards related to corporate governance, or even Securities and Exchange laws if such companies are publicly traded. In particular, there is a risk that a company's board of directors or executive officers may not be fulfilling fiduciary obligations to their company's shareholders if such persons engage in the delegation of authority to Agents without appropriate parameters or supervision. There is a possibility that a publicly traded corporation may violate its Securities and Exchange Act registration statement and reporting obligations due to such disclosure deficiencies – and ultimately be subject to stockholder derivative litigation – if it were to expose its workforce to risk from potential financial, operational or reputational misconduct without any discussions of its delegating responsibilities over such activities.

Moreover, it has yet to be determined whether legislation that establishes liability for negligence would apply to delegated Agent actions, both legally or practically. There is considerable uncertainty as to how liability is determined for Agent actions that cause civil harm or property damage. In this regard, determining liability for such activities performed by Agents could depend upon whether the Agent is classified as a product, or as a service. It is also uncertain whether tortious activity performed by a Non-Agent would include negligent or defective product actions that break applicable tort laws.

Statutes like the Right to Privacy Act and similar provisions in the United States, as well as privacy regulations and cyber security guidelines could also apply.

3.9. Future Trends in Agentic AI

While the research around agentic AI is nascent, we can draw inferences from current trends in AI. The specific applications that involve agentic AI will depend on how capable future AI systems become, how integrated with other technologies these systems become, and the specific characteristics of the future work environment. There are a number of other possible future trends and questions to ask around future trends in agentic AI. The first is: What future advancements in AI capabilities can we expect? Much research seeks to build general agents that can operate across many categories of tasks.

Fig : AI's Strategic Decision Making

The progress made in the specific capabilities around reaction time speed, memory, and planning for existing agents, along with the growing interest and funding behind AGI from academia and investors, suggest that we can expect at some point much more capable, generalizable AI agents. The other specific focus of past works on AGI suggest that these advancements might entail systems with the ability to understand agents beyond surface measurements, use multi-modality, explain versus imitate, and use common sense reasoning. These capabilities will improve the types of tasks these systems can take on independently, as well as the human workflows and decision making processes they handle. At the moment, the capabilities of existing AI tools are not in a stage where deployed digital agents will support automation enabled decision making in the vast majority of areas within a company.

One additional major trend to consider is the growing integration of AI with other emerging technologies. There are many areas in which AI agents are being combined with other emerging technologies. This includes robotics and IoTs. Other areas include the use of AI in computer graphics, extended reality platforms, and generative deep fakes, which will expand the impact AI agents have in the design, marketing, and influence domains. Another possible subarea of research interest is the area of increasing convergence around federated and distributed AI, and regulating the use of AI systems, from responsible and ethical AI to the area of AI to help with malicious uses of AI, security, and forensics. As mentioned before, research in these areas is currently limited.

3.9.1. Advancements in AI Capabilities

Understanding and defining Agentic AI is a complex and difficult task because it is based on broad, complex ideas, and it is rooted in an even larger and complicated topic — AI technology. Recent advances in AI technologies of different types, and particularly the convergence of those advances into integrated systems and solutions, have fueled the current explosion of excitement, interest and innovation in Agentic AI initiatives. Some Agentic AI initiatives are surface enhancements to existing technology platforms yet many initiatives are fundamentally new, exploring the new frontiers of AI technology and harnessing the inherent capabilities of potentially autonomous self-evolving digital agents. Advances in AI are described in detail in many other works and we will only briefly summarize some developments related to the capabilities of Agentic AI in this section.

In its broadest expanses, AI is defined as the study and development of systems that display intelligent behavior. Over the years, however, the development of AI technology has moved away from requiring human-like intelligence, reasoning, and understanding toward greater application of experience to solve specific problems. In short, AI for a large portion of the recent past has become an extremely effective tool for what is

described as bound rationality – rather than generating and then exploring all possible solutions to a problem, the algorithm is designed to explore a tractable, small portion of the total outcome solution space, and then exploit what it determines is the best outcome among the small solutions of the total space.

3.9.2. Integration with Emerging Technologies

In this essay, we explore how agentic AI will influence the introduction and integration of emerging technologies, including the Internet of Things, digital twins, quantum computing, immersive and ambient interfaces, the metaverse, edge computing, and decentralized computing. These technologies serve as novel substrates for agentic AI enablement. Their integration is built over the fundamentals of cloud-native business architectures created during the prior digital transformation wave. Although the agents will rely on cloud services running state-of-the-art large language models and computer vision, newer AI computing substrates, such as storage-class memory and embedded accelerators, will be necessary to perform real-time actions while preserving end-user privacy. Additionally, both edge and decentralized computing environments will support lower-stakes environments, helping users build their AI agents at low-cost and low-risk conditions.

The IoT and digital twins will drive a significant increase in the number of agentic AI agents acting autonomously on behalf of businesses, families, and individuals. Emerging technology ecosystems where digital twins reproduce key operational aspects of physical or abstract environments supported by embedded IoT devices will require considerable automation to thrive. Agentic AI accelerates this automation by authoring the workflows and processes that control the digital twins and by executing atmospherics: the remote actions done through embedded IoT devices that affect the physical counterparts of the digital twins. Compute clouds will provide additional AI capabilities to small and local embedded devices, such as increased natural language understanding capabilities when processing user voice commands. Edge clouds will facilitate the low-latency needs for the sensor fusion and automatic multimedia content generation capabilities necessary to experience ambient intelligent environments.

3.9.3. The Role of AI in Future Work Environments

The future of work is poised to undergo a significant transformation precipitated by the continued advancement of AI technologies, including agentic AI. Businesses will become increasingly dependent on emergent, more basic AI capabilities to catalyze operational efficiency and success, even in the face of potential ethical concerns. The future digital workforce will inevitably include both non-human agents and human

workers, necessitating a thoughtful consideration of the impact of these technologies on existing roles, capabilities and interactive arrangements. The coming together of agentic AIs and humans in the workplace also raises interesting questions about the nature of collaboration with intelligent entities that work independently but in support of human goals.

In the foreseeable future, as AI systems steadily grow more capable and integrated into workplace workflows, highly educated and skilled professionals will make extensive use of Intelligent Augmenters in their everyday work projects. But we can also expect to see agentic AI embedded in the everyday tasks of many other workplace participants, from teachers in classrooms to retail sales associates, as an Intelligent Assistant that can take care of repetitive higher-level tasks, handle menial tasks ordinarily delegated to ordinary human workers, or even engage customers and clients in low-risk situations. Such agentic tools will free up human workers to focus on higher-level collaborative tasks that truly require human social creativity. Over time, as more practical agents are allowed to perform simple work management for both mundane and skilled tasks, a better understanding of the benefits and risks of human interaction with non-human agents will begin to emerge. This foundation will make it practical to experiment intentionally with methods of collaboration and supervision, first with the safer Intelligent Assistants, and later with more intelligent agents.

3.10. Case Studies

Detailed in this section are several use cases that highlight some of the benefits and successes that might be found in implementing Agentic AI, as well as some of the challenges and issues that can arise in doing so. This section is by no means comprehensive in providing case studies; there is a wide variety of implementations of each of the various implementation strategies outlined above described in an entire universe of publications. However, as an introduction, we picture the general landscape of meaningful implementations of autonomous digital agents, particularly those within business settings and functions, by providing illustrative examples.

3.10.1. Successful Implementations of Autonomous Agents

Business and research interest in implementing Agentic AI is increasing, as many organizations have begun to implement simple task-oriented agents – if-framed agents and hierarchical behavioral model agents. For example, there are significant numbers of early implementations in marketing, advertising, and information dissemination. These include companies that use automated bots and independent accounts to engage in misleading, deceptive, or malicious activity on social media platforms; such bots can

reach a far greater audience than can individual humans and can do so while spreading widely and rapidly misinformation. Examples include candidates for political office using chatbots; national news agencies using automated programs to generate reports; and businesses and organizations using bots to respond to customer queries on social media platforms.

3.10.1. Successful Implementations of Autonomous Agents

Over the years, early adopters have developed hidden tech gems: digital agents that people refuse to discard, and the reasons reveal a world of possibilities: a fuel reconciliation bot built to check the cost per mile of every flight for an airline. Bots specialized in finding, optimizing, and monitoring thresholds of real-time weather data and providing expedient reactions to the party planning a wedding or other outdoor events. Bots that monitored midnight petitions filed by federally registered corporations to determine who was filing for bankruptcy, and then messaging timeshares about putting together a "low price" group to buy those bankrupt properties, in the market for pennies on the dollar. Bots that monitored filings for the same purpose, but were then able to donate 10% of the proceeds to worthy causes instead of fattening the brokers' wallets as happened in the timeshare example. Bots that filed documents to follow the rules of the court with the same fervor as teenagers who are told they might lose their driving privileges if they do not smile in their ID photos.

These bots have been implemented as side machines doing no harm and providing monstrous return on investment, and they are too cheap to infringe on social norms enough to introduce systemic risks. It has been summarized that "the best processes to automate and/or augment via Agents are the repetitive, boring, and mundane tasks. These tasks, when completed by humans, do not utilize humans at their full potential, but instead result in disengagement, boredom, and importantly, mistakes, which in many cases can be costly. Furthermore, these tasks often rely on data from systems that have rigid rules of engagement and offer limited alternatives when it comes to automating the flow of information. These are the perfect scenarios to apply Agent technology. Perfect, but also tricky. Tricky because only once do you want to Roll It Out Into Vendor Production."

3.10.2. Lessons Learned from Failures

Many AI systems have become increasingly complex over time, as such systems remain an area of active technological innovation and provide support for many areas of life. Yet they still remain often highly specialized and incapable of generalizing to new situations. This essential nature of AI capabilities has resulted in many projects failing

to reach an expected level of expressiveness, correctness, safety, security, or utility. As a result, lessons learned from failures may be of almost as much interest as classic implementations of autonomous agent and AI systems. Understanding what went wrong, how the situation was managed, and what such agents do badly can capture our interest into how the dangers of any future such agents may be minimized. This is especially true as society and especially people in control of high-stakes decisions and actions rely more closely on autonomous and semi-autonomous digital agents. The chances for failure and harm become higher as the effects of such agents on every facet of human existence increase. Design decisions about agent functions or their tasking and skills may have not been anticipated correctly, resulting in significant behavioral gaps. Lack of sufficient training data on which to train such agents, or an agent dealing with tasks outside its anticipated regions of competency, could lead to unexpected behaviors. Frustrated by such behaviors, agents may be abandoned or overly limited in function, blocking potential benefits. The socialization and accountability aspects of agents cannot be ignored. It is only recently that researchers and developers have become both significantly aware and concerned enough to study these issues and defects in sufficient detail. For these reasons, this section will discuss major projects that experienced important defects or deficiencies. Far from being avoided or hidden, these problem projects are documented and widely reported, and could make important contributions to addressing both required and desired progress in agentization, not only in higher-level intelligent approaches but in their implementations in AI implementations. Less ironically, we are at present in a position to learn from these negative experiences.

3.11. Conclusion

Our primary goal in this work is to provide a clear and coherent conceptual framework that can catalyze further reflection about these questions and motivate research aimed at illuminating the dark corners and hidden chasms. We want to avoid the traps that lie ahead as humanity proceeds along its grand intellectual journey, exploring the implications of the creation of increasingly capable agentic AI. Our observation is that we already have a rich conceptual and terminological pool upon which to draw. Rather than just inventing more words and new concepts because we can, we should seek to consolidate our understanding and the models we have, deepen and explore them, and build bridges between disciplines, activating collective intelligence.

Our model provides a horizontal axis describing increasing levels of agency and a vertical axis describing increasing levels of extranormality of intelligence. This simple two-dimensional structure is able to capture a wealth of concepts ranging from current philosophical and scientific inquiries regarding the concept of person, to more technical approaches that stem from the foundation of intelligence technology, such as Informed

Agency and Policy Search. Our Unified Conceptual Model brings together Multiple-In Single-Out architectures focusing on a single entity, paths probing into extranormal spaces, and the special case of the Utility Agent modeling the applications of existing mainstream AIs considering their special nature.

The framework helps place current advanced AI technologies and their applications in the landscape of models surrounding the two axes. It also allows us to identify paths leading past common modeling assumptions and guidelines, often left implicit, and proposing new ones. The concept and perspective proposed by the Unified Conceptual Model are but one thread of a complex tapestry that invites contributions from the reader, as it stitches together the frontiers of the challenges faced when exploring the question of agentic AI.€

References:

S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed., Pearson, 2020.

P. Stone et al., "Artificial intelligence and life in 2030," One Hundred Year Study on Artificial Intelligence, Stanford Univ., 2016.

T. Bryson, J. Cave, and J. R. Cave, "Agentic computing and autonomy: Challenges for AI governance," IEEE Technol. Soc. Mag., vol. 39, no. 1, pp. 56–64, 2020.

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, 1-29.

Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous Intelligence for Complex Goals–A Comprehensive Survey. IEEE Access.

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenge. arXiv preprint arXiv:2505.10468.