ETHICAL CONCERNS AND MASS AGENTIC AI ADOPTION

Sonja Jovanović[1]*,
[0009-0009-3538-8096]
Lazar Dražeta[2],
[0000-0001-5246-6468]
Aleksandar Petrović[2],
[0000-0003-3324-3909]
Nebojša Bačanin Džakula[2],
[0000-0002-2062-924X]
Miodrag Živković[2]
[0000-0002-4351-068X]

[1]Ernst & Young,
Belgrade, Serbia

[2]Singidunum University,
Belgrade, Serbia

Abstract:
The arrival of agentic artificial intelligence (AI) marks a significant jump in the evolution of intelligent systems. Unlike AI so far, which primarily operates within predefined parameters, agentic AI possesses the capability to autonomously perceive, reason, act, and learn from its environment. This new paradigm enables AI agents to solve complex, multi–step problems and interact with various tools and data sources independently. Rather than executing commands or routines set by humans, Agentic AI systems are intended to adjust their strategies and explore their environments as needed. Agentic AI's sophisticated reasoning and iterative planning abilities are settled to transform numerous industries. From optimizing supply chains and personalizing customer service to enhancing healthcare and automating complex tasks, the potential applications are vast and varied. However, with these advancements come challenges, including the need to address biases, ensure ethical use, and establish robust regulatory frameworks. This paper explores the foundational principles of Agentic AI, its current and potential applications, and the ethical considerations that must guide its development and deployment. By examining these aspects, the aim is to provide a good understanding of how Agentic AI can be used to drive innovation while mitigating associated risks.

Keywords:
Agentic AI, Evolution, Autonomous Decision Making, Use Cases.

INTRODUCTION

Agentic artificial intelligence (AI) is an autonomous or semi–autonomous software entity that can process data, form decisions and take actions in the digital and physical world. It can process information and perform tasks, rather than just generating answers like many of the chatbots seen in recent years.

AI chatbots use generative AI to provide responses. It is important to highlight that chatbot response will be original, but highly dependent on already known or most frequently repeated information. When a person asks a question, the chatbot utilizes natural language processing to respond. Agentic AI goes beyond [1]. It stands out due to its autonomy and ability to independently pursue specific goals while learning from feedback. Agentic AI continuously monitors and processes data inputs, as well as inputs from the users.

Correspondence:
Sonja Jovanović

e-mail:
sonja.jovanovic@rs.ey.com

DOI: 10.15308/Sinteza-2025-369-375

Based on those inputs, it autonomously makes informed decisions and takes actions, if needed. It also learns from its experiences and mistakes. In essence, Agentic AI can autonomously perform complex tasks in real time [2]. The adoption of generative AI has surpassed all the entertainment platforms in terms of adoption time, as illustrated in Figure 1. The speed at which it was adopted is immense considering that the entertainment industry is of much higher interest to the general public. However, the impact of AI and visions of a cyberpunk future have prevailed against the interests of the average person.

When examining the differences between generative AI and agentic AI, three main points are notable: it focuses on decision–making rather than content creation, it operates independently, targeting specific goals such as increasing sales, enhancing customer satisfaction, or improving employee experiences, without needing human prompts, and agentic AI can perform complex tasks and independently search through various databases and data sources.

Mass application of agentic AI is the next step, as its use cases are vast. This work reviews its numerous applications and provides a perspective on the ethical dilemmas associated with it. Without considering the possible job shortages that such systems would produce, healthcare concerns are first to be raised [4]. There must exist a standard before approving that the AI in such an environment is safe for human use. Therefore the goal of this work is to provide a connection between the use cases and risks that they bring with their use.

2. AGENTIC AI USE CASES

Agentic AI can boost productivity and efficiency for security, financial operations, software development, customer engagement, digital marketing, knowledge acquisition, and more [5]. As Agentic AI can self–correct and improve over time, it is expected to accelerate, optimise and boost productivity and system performance.

Predictions from Oracle and Gartner are saying that Agentic AI will soon find its way into 33% of enterprise software applications, replacing 20% of human interactions and making 15% of day–to–day workplace decisions by 2028 [6]. In addition, the global market for Agentic AI is projected to reach $120 billion by 2030, driven by ever–increasing investments in AI technologies and their applications across various industries [7].

Autonomous vehicles are well–known examples of agentic AI today. These vehicles use multimodal models to process data from sensors like cameras, light detection and ranging (LiDAR), a technology that measures distance with laser beams, and radar. This helps them understand their surroundings, recognise objects, and make decisions for safe navigation. They use predictive algorithms to respond faster than humans, preventing dangerous situations. By processing sensor data and using deep learning algorithms, autonomous vehicles can determine the best route, reduce travel time, and identify obstacles in real time. Waymo's driverless cars, originating from Google's self–driving car project is the world's earliest autonomous ride–hailing service and a great example of Agentic AI in action [8].

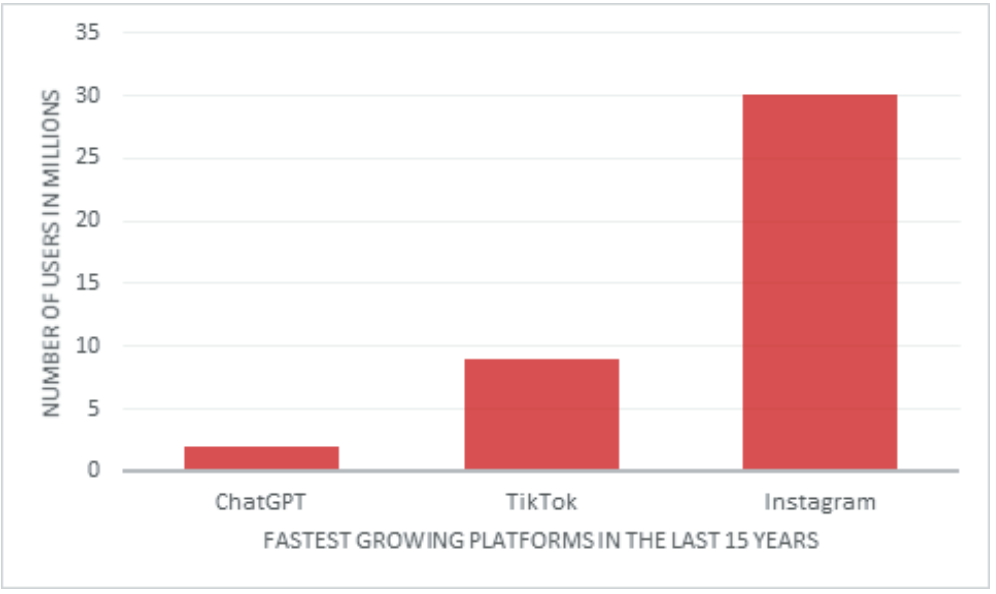

Figure 1. Speed of adoption of revolutionary technological platforms [3]

Agentic AI systems, with their advanced decision–making and execution abilities, are excellent platforms for experimentation and innovation. Multi–agent AI models can quickly scan and analyse large research spaces, like scientific articles and databases, much faster than human researchers. SciAgents, developed by MIT, are robot scientists to create research plans and a critic agent to review and improve them [9]. As an example, these two AI agents identified a new biomaterial made from silk and dandelion pigments, which has better mechanical and optical properties and requires less energy to produce.

An Agentic customer service agent can predict if a delivery will be late, notify the customer about the delay, and offer a discount to ease their disappointment. Agentic AI can search through thousands of databases and apps to solve customer queries and complaints [10]. These chatbots learn from each interaction and suggest actions for human agents. They can also check its content for accuracy and compliance and recommend improvements to the customer knowledge base.

For sales agents, the essential task of identifying and nurturing sales leads can often be overwhelmed by numerous emails, paperwork, and other routine but necessary administrative duties. AI systems designed for sales can significantly reduce the time spent on these activities. For instance, Salesforce has introduced its Agent Force service development rep [11] to support human sales teams. Using large language models (LLM), this AI can understand customer messages, suggest follow–up actions, schedule meetings, answer inquiries, and create responses that match the company's brand voice. Additionally, the agent force sales coach offers personalized feedback to sales agents and provides learning opportunities through virtual role–play sessions.

The ability of agentic AI systems to adapt to various environments, understand human emotions, and demonstrate empathy makes them well–suited for non–routine, soft–skills tasks in fields like healthcare and caregiving. Hippocratic AI [12], a healthcare company based in California, has developed a range of AI agents specialized in different aspects of healthcare and social support. Among these agents is Sarah – who is known for her warmth and understanding in assisting with daily living activities. Sarah can engage with patients about their day, manage meal plans and transportation, and remind them to take their medication. Another agent, Judy, supports patients with pre–operative procedures by providing reminders about arrival times, locations, and instructions on fasting or medication adjustments.

In the gaming industry, Agentic AI has a lot of things to offer. Enhancing the behaviour of bots in video games is one of them. Non–player characters (NPC) are characters in video games that are not controlled by the human player – like enemies, allies, merchants, quest givers, etc. While traditional NPCs follow pre–programmed scripts or simple learning algorithms, agentic AI enables NPCs to exhibit more complex and adaptive behaviours. By using the reinforcement learning technique, NPCs can learn from their environment and improve their actions over time, based on rewards and penalties [13]. For example, NPCs can engage in more natural and varied conversations, responding to the player's choices, or if a human player acts aggressively, NPCs might become more hostile or wary.

There is significant potential for using Agentic AI in human resource management. For example, in talent management, AI agents can analyse employee performance data, identify skill gaps, and recommend personalized development plans [14]. This approach ensures that employees receive targeted training and opportunities for career growth. Another valuable application is in employee engagement and retention. Agentic AI can monitor employee satisfaction through sentiment analysis of feedback and interactions. It can also suggest interventions to enhance employee experience and engagement, that result in reducing turnover.

The adoption of AI in some countries is high according to the IBM AI adoption index [15]. This index shows the percentage of enterprises that actively deploy technologies reliant on AI. The data is self–reported by the companies included in the survey. The country with the highest AI adoption index is India, amounting to almost 60% usage. The country with the lowest adoption rate was Australia, however, this should not be considered generally as bad, since most of the countries do not even make it on this list.

There are a lot of talks about AI, but what will really separate companies from their competitors, when it comes to AI agenda – is reliable, legal, and proprietary data at scale. According to EY research, 83% of senior leaders recognise there is a gap in their capabilities and believe that their AI adoption would accelerate if they had stronger data infrastructure. Furthermore, 67% of senior leaders also admit their lack of infrastructure is actively holding back AI adoption [16].

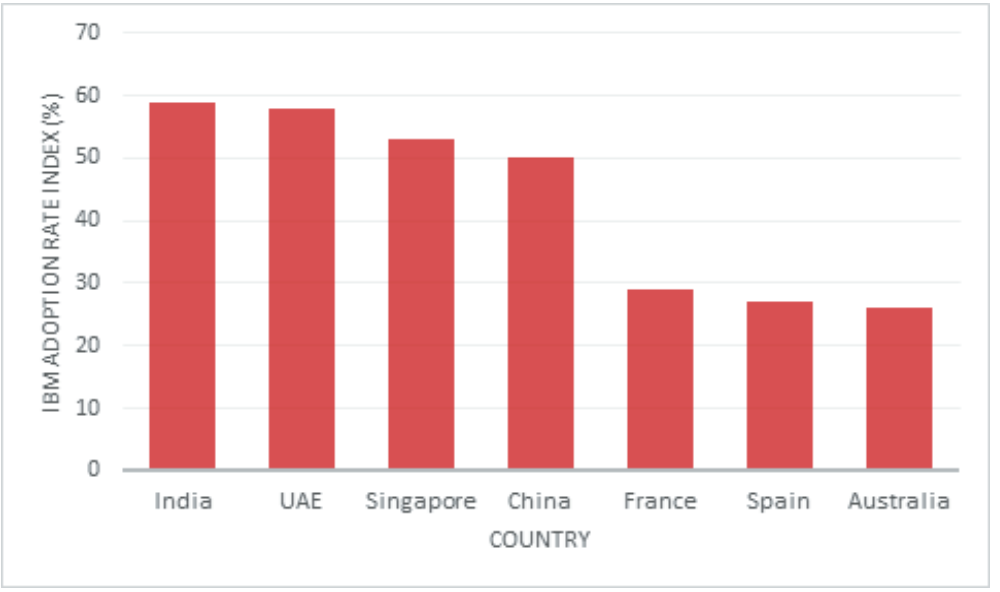

Figure 2. IBM AI adoption rate index [15]

This concept emphasizes that the quality of output from an AI system is directly dependent on the quality of the input data it receives. For Agentic AI, which relies on vast amounts of data to learn, reason, and make decisions autonomously, ensuring high–quality, accurate, and unbiased input data is crucial. Poor quality data can lead to flawed reasoning, biased outcomes, and ineffective actions, undermining the potential benefits of such advanced AI systems [17]. For this reason, Agentic AI may not be perceived as completely trustworthy. Similar to how human teams can struggle with unclear or poorly communicated goals, Agentic AI systems can also fail if their objectives are not clearly defined.

3. RISK MANAGEMENT

As AI models become more advanced and agentic, they may also become less predictable and sometimes more ethically ambiguous. To address these risks, organizations should establish guidelines, risk management frameworks, and governance protocols. These measures will help ensure that Agentic systems prioritise legality, data privacy, security, human rights, and corporate accountability.

It is crucial to develop strategies for securely integrating this technology with existing system architecture. Additionally, it is important to create fail–safe mechanisms to prevent systemic failures. Regular behavioural monitoring, observation, and system analysis should also be conducted.

The most important feature is the presence of a human–in–the–loop (HITL), ensuring that employees are strategically positioned to monitor and oversee agentic operations. Information security teams considering AI agent integration should review the entire workflow from an automation–only perspective and then determine how to reintegrate human personnel into the agentic workflow. This requires education and training to ensure the workforce understands AI capabilities, benefits, and limitations.

While the risks and vulnerabilities are addressed, there are also costs associated with deploying and securing AI agents in enterprise environments. Additionally, there are unique governance challenges related to security assurance, regulation, legal accountability, data equity, and interoperability.

With great power comes great responsibility. As AI initiatives yield benefits, senior leaders must address the ethical implications and risks. Interest in responsible AI among senior leaders has increased over the past year, 61% vs. 53% six months ago [16]. Respondents also indicate that this interest will continue to grow over the next year. Consequently, more organizations are dedicating time to training employees in responsible AI and increasing transparency with customers about AI usage.

The variety of fields affected by AI also poses another threat to safety. This technology is making its way into critical fields like EEG predictions [18], intrusion detection [19], crop yield [20], gold price prediction [21], and some not–so–critical but nonetheless very important like software defect prediction [22], fake news detection [23], and machine learning optimization [24].

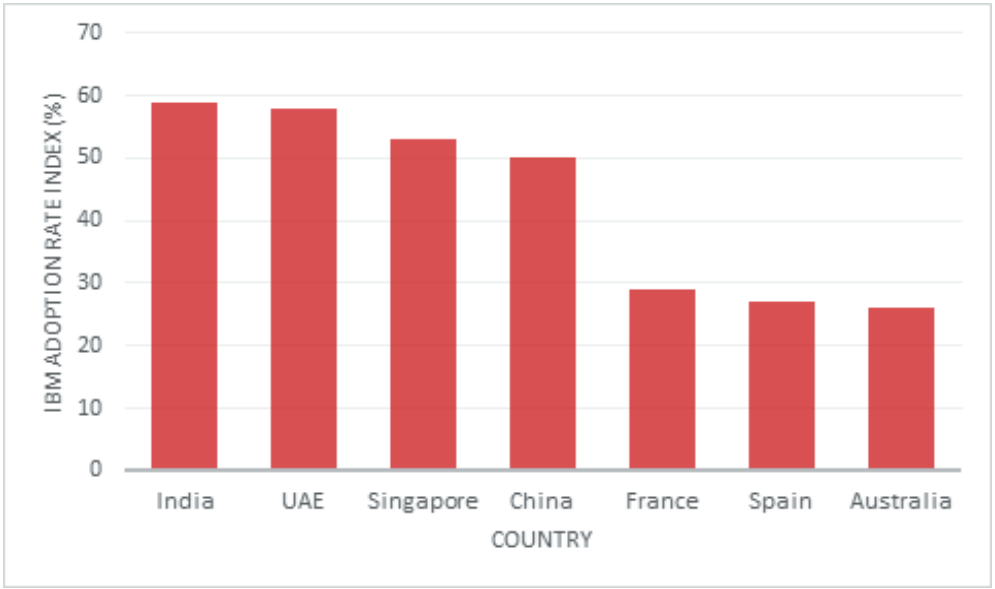

Figure 3. Hallucination rates of the popular generative agents [25]

The risks of AI use are depicted in Figure 3. It is shown how much of the data was hallucinated by generative agents. This is one of the most important topics regarding generative agents. The phenomenon of hallucination in generative agents can be described as making up information. This is a tendency of such agents when there is no usable data and no information in their knowledge base. The importance of this phenomenon is for the real–world use of such agents. They cannot be applied in critical fields until the rate of hallucinations is brought down to zero.

4. CONCLUSION

Agentic AI represents a significant advancement in AI, with the potential to transform various sectors. These systems, characterized by their ability to adapt, reason, and exhibit empathy, offer promising solutions for non–routine, soft–skills tasks. Agentic AI systems are still in the early stages of development. Although they have advanced reasoning and execution abilities, they do not eliminate traditional workforce management challenges – they transform them. Similar to managing human teams, managers must consider team composition and role selection and set appropriate goals to ensure the success of Agentic AI or hybrid teams. Additionally, managers need to determine when Agentic AI systems can be trusted to make decisions and when human intervention is necessary.

To utilise the full potential of Agentic AI, it is essential to establish comprehensive guidelines, risk management frameworks, and governance protocols. Ensuring the presence of a HITL is crucial for monitoring and overseeing AI operations. Upskilling, reskilling, education, and training programs are necessary to equip the workforce with a functional understanding of AI capabilities, benefits, and limitations. Further advancements in AI should bring full autonomy, which raises the question if it should be researched at all.

The lack of regulation is still a massive issue and the biggest threat from AI. This is an experimental technology and should be treated as such, which requires some security precautions. Current awareness on this topic is low, and that is why this paper is important as it serves to provide a review of the current state of agentic AI and the risks of its use.

REFERENCES

[1] D. B. Acharya, K. Kuppan and B. Divya, "Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey," in *IEEE Access*, vol. 13, pp. 18912–18936, 2025, doi: 10.1109/ACCESS.2025.3532853.

[2] B. L. Aylak, "SustAI–SCM: Intelligent Supply Chain Process Automation with Agentic AI for Sustainability and Cost Efficiency," *Sustainability*, vol. 17, no. 6, p. 2453, 2025.

[3] K. Hu, "ChatGPT sets record for fastest–growing user base – analyst note," Reuters, 2023 February 2023. [Online]. Available: https://www.reuters.com/technology/chatgpt–sets–record–fastest–growing–user–base–analyst–note–2023–02–01/. [Accessed 15 March 2025].

[4] T. A. Bach, J. K. Kristiansen, A. Babic and A. Jacovi, "Unpacking Human–AI Interaction in Safety–Critical Industries: A Systematic Literature Review," in *IEEE Access*, vol. 12, pp. 106385–106414, 2024, doi: 10.1109/ACCESS.2024.3437190.

[5] S. Samdani, K. Paul, and F. Saldanha, "Agentic AI in the Age of Hyper–Automation," *World Journal of Advanced Engineering Technology and Sciences*, vol. 08, no. 01, pp. 416–427, 2023.

[6] D. Sun, "Capitalize on the AI Agent Opportunity," Gartner, 27 February 2025. [Online]. Available: https://www.gartner.com/en/articles/ai–agents. [Accessed 10 March 2025].

[7] L. Craig, "What is agentic AI? Complete guide," TechTarget, September 2024. [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/agentic–AI. [Accessed 10 March 2025].

[8] B. Xia, J. Zhou, F. Kong, Y. You, J. Yang, and L. Lin, "Enhancing 3D object detection through multi–modal fusion for cooperative perception," *Alexandria Engineering Journal,* vol. 104, pp. 46–55, 2024.

[9] A. Ghafarollahi, and M. J. Buehler, "SciAgents: Automating Scientific Discovery Through Bioinspired Multi–Agent Intelligent Graph Reasoning," *Advanced Materials*, p. 2413523, 2024.

[10] D. Leocádio, L.Guedes, J. Oliveira, J. Reis, and N. Melão, "Customer service with AI–powered human–robot collaboration (HRC): A literature review," *Procedia Computer Science*, vol. 232, pp. 1222–1232, 2024.

[11] Salesforce, "Meet Einstein SDR and Einstein Sales Coach: Two New Autonomous AI Sales Agents to Scale Your Sales Team," Salesforce, 22 August 2024. [Online]. Available: https://www.salesforce.com/news/stories/einstein–sales–agents–announcement/. [Accessed 10 March 2025].

[12] H. AI, "Hippocratic AI: Safety Focused Generative AI for Healthcare," Hippocratic AI, 2025. [Online]. Available: https://www.hippocraticai.com/. [Accessed 10 March 2025].

[13] A. Maciá–Lillo, A. Jimeno–Morenilla, H. Mora and E. Duta, "Hybrid Architecture for AI–Based RTS Games," in *IEEE Transactions on Games*, doi: 10.1109/TG.2025.3533949.

[14] M. Rožman, D. Oreški, and P. Tominc, "Integrating artificial intelligence into a talent management model to increase the work engagement and performance of enterprises," *Frontiers in psychology*, vol. 13, p. 1014434, 2022.

[15] E. Shein, "IBM: While Enterprise Adoption of Artificial Intelligence Increases, Barriers are Limiting Its Usage," TechRepublic, 12 January 2024. [Online]. Available: https://www.techrepublic.com/article/ibm–global–ai–adoption–index/. [Accessed 10 March 2025].

[16] L. McWilliams, "EY research: Artificial intelligence investments set to remain strong in 2025, but senior leaders recognize emerging risks," EY, 10 Dec 2024. [Online]. Available: https://www.ey.com/en_us/newsroom/2024/12/ey–research–artificial–intelligence–investments–set–to–remain–strong–in–2025–but–senior–leaders–recognize–emerging–risks?utm_source=chatgpt.com. [Accessed 10 March 2025].

[17] A. Majeed and S. O. Hwang, "When Poor–Quality Data Meet Anonymization Models: Threats and Countermeasures," in *IEEE Access*, vol. 13, pp. 49457–49475, 2025, doi: 10.1109/ACCESS.2025.3552412.

[18] N. Bacanin, L. Jovanovic, A. Toskovic, M. Zivkovic, A. Petrovic, and M. Antonijevic, "Anomalous EEG Signal Time Series Classification Using Modified Metaheuristic Optimized RNN," in *International Conference on Communication and Intelligent Systems}*, Springer, 2023, pp. 291–304.

[19] A. Petrovic, L. Jovanovic, M. Antonijevic, N. Bacanin, M. Zivkovic, and J. Kaljevic, "Natural Language Processing of HTTP Content for Insider Threat Detection Optimized by Modified Metaheuristic," in *International Conference on Innovations in Cybersecurity and Data Science Proceedings of ICICDS*, Springer, 2024, pp. 299–314.

[20] N N. Bacanin et al., "Crop Yield Forecasting Based on Echo State Network Tuned by Crayfish Optimization Algorithm," *2024 IEEE International Conference on Contemporary Computing and Communications (InC4)*, Bangalore, India, 2024, pp. 1–6, doi: 10.1109/InC460750.2024.10649266.

[21] S. Golubovic, A. Petrovic, A. Bozovic, M. Antonijevic, M. Zivkovic, and N. Bacanin, "Gold price forecast using variational mode decomposition–aided long short–term model tuned by modified whale optimization algorithm," in *International Conference on Data Intelligence and Cognitive Informatics,* Springer, 2023, pp. 69–83.

[22] A. Petrovic, L. Jovanovic, N. Bacanin, M. Antonijevic, N. Savanovic, M. Zivkovic, M. Milovanovic, and V. Gajic, “Exploring metaheuristic optimized machine learning for software defect detection on natural language and classical datasets,” *Mathematics*, vol. }, no. 18, p. 2918, 2024.

[23] A. Petrovic, J. Perisic, L. Jovanovic, M. Zivkovic, M. Antonijevic and N. Bacanin, "Natural Language Processing Approach for Fake News Detection Using Metaheuristics Optimized Extreme Gradient Boosting," *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*, Gwalior, India, 2024, pp. 252–257, doi: 10.1109/AIC61668.2024.10731062.

[24] A. Vesic, M. Marjanovic, A. Petrovic, I. Strumberger, E. Tuba and T. Bezdan, "Optimizing Extreme Learning Machine by Animal Migration Optimization," *2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia, 2022, pp. 261–266, doi: 10.1109/ZINC55034.2022.9840711.

[25] M. Chelli, J. Descamps, V. Lavoué, C. Trojani, M. Azar, M. Deckert, J. L. Raynier, G. Clowez, P. Boileau, and C. Ruetsch–Chelli, “Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: comparative analysis,” *Journal of medical Internet research*, vol. 26, p. e53164, 2024.