MEMÓRIAS
DA
ACADEMIA DAS CIÊNCIAS
DE
LISBOA

CLASSE DE CIÊNCIAS

The Impact of Artificial Intelligence in Science

ARLINDO L. OLIVEIRA

ACADEMIA DAS CIÊNCIAS
DE LISBOA

LISBOA • 2025

Título: The Impact of Artificial Intelligence in Science
Edição: Academia das Ciências de Lisboa
Data de edição: 2025
DOI: https://doi.org/10.58164/4xhs-r954

The Impact of Artificial Intelligence in Science

Arlindo L. Oliveira[1]

ABSTRACT

Generative artificial intelligence has significantly impacted scientific research by enabling new methods for data analysis, hypothesis generation, and experimental design. These technologies have accelerated discoveries in fields such as drug development, materials science, and climate modeling by automating the generation of novel solutions and simulations. By enhancing data-driven insights, generative AI has also improved the accuracy and efficiency of predictive models, opening new avenues for interdisciplinary research. However, the use of artificial intelligence raises ethical concerns regarding data integrity, reproducibility, fake results, and the potential for bias, highlighting the need for responsible and transparent implementation in scientific processes and workflows. This article addresses several of these issues, using as the starting point the SAPEA evidence review report on the successful and timely uptake of artificial intelligence in science.

INTRODUCTION

The advent of artificial intelligence (AI) marks a profound shift in the way scientific research is conceived, conducted, and communicated. Once confined to theoretical exercises, toy problems and a limited range of applications, AI has evolved into a general-purpose technology capable of transforming entire domains of human knowledge. Among the most promising and disruptive developments is generative AI, a class of models and techniques that can autonomously generate new contents, create hypotheses, analyse complex datasets, and design novel experimental frameworks. These advances are not mere enhancements of existing tools; they represent a restructuring of the epistemic and methodological foundations of science.

AI is not a monolitic area, but rather a constellation of technologies ranging from machine learning and deep neural networks to symbolic reasoning, discrete optimization methods and multimodal generative systems. Among these, large-scale language models and reinforcement learning agents have found applications in drug discovery, protein structure prediction, materials design, climate

[1] Academia das Ciências de Lisboa e Instituto Superior Técnico/INESC-ID

modelling, and more. AlphaFold's success in predicting protein folding, AI-driven identification of new antibiotics, and advanced weather forecasting systems are only a few examples of the ability of AI to tackle long-standing scientific challenges with unprecedented efficiency and scale.

Recognising the transformative potential of these technologies, the European Commission has sought to promote a responsible and timely uptake of AI across scientific domains. In its March 2024 report on the successful adoption of AI in science — supported by the Science Advice for Policy by European Academies (SAPEA) — the Group of Chief Scientific Advisors provided an in-depth assessment of how AI is reshaping the European research ecosystem (SAPEA, 2024). The report identifies both the considerable opportunities and the complex risks involved in this technological transition.

At the core of AI's value in science lies its ability to accelerate discovery. Through automation of data analysis, generation of testable hypotheses from literature, and simulation of complex systems, AI enables researchers to operate beyond the constraints of conventional analytical methods. It fosters interdisciplinary collaboration and facilitates the integration of vast and heterogeneous datasets across fields, from quantum physics to digital humanities. AI-driven experimental design and autonomous laboratories are already redefining the workflows of materials science and molecular biology.

However, the scientific community must navigate these possibilities with caution. AI's opacity — the so-called "black box" problem — poses a serious threat to the principles of transparency, reproducibility, and interpretability that are foundational to the scientific method. Moreover, AI's performance is only as robust as the data it is trained on; biased, incomplete, or non-representative datasets can lead to flawed models that amplify existing inequalities or propagate misinformation. The increasing reliance on proprietary models developed by private tech companies also raises concerns regarding access, reproducibility, and the monopolisation of knowledge.

From a structural standpoint, AI is impacting not only the methods used in the scientific practice but also its social fabric and governance. The integration of AI tools into research practices changes the nature of scientific labour, raising questions about training, career paths, and mental well-being. While AI can offload repetitive tasks and free researchers for more creative inquiry, it also demands new skills in data management, computational thinking, and ethical reasoning. The SAPEA report underscores the need for broad-based training initiatives that span disciplines and educational levels, alongside robust infrastructures for AI experimentation and evaluation.

Furthermore, the environmental impact of training large-scale models, the potential misuse of AI in scientific fraud or bioweapons development, and the ethical dilemmas posed by autonomous decision-making systems all call for careful governance. The European Union, through initiatives such as the Artificial Intelligence Act has begun to establish a framework that balances innovation with ethical oversight. Yet is is clear that policy must evolve rapidly to keep pace with technological change, a requirement that represents a significant challenge.

The SAPEA evidence report proposes several structural interventions to support this transition: the creation of a distributed European institute for AI in science (EDIRAS), the establishment of an AI Science Council, and the development of green and trustworthy AI tools tailored to scientific research. These proposals reflect a strategic vision in which Europe leads not just in technological capacity, but in setting global standards for responsible AI in science.

This article examines how generative AI and related technologies are impacting scientific inquiry today — technically, institutionally, and ethically. Drawing on the insights from the SAPEA evidence report and recent research, it aims to illuminate the pathways through which AI is becoming a catalyst for innovation, while also highlighting the imperatives for transparency, inclusiveness, and accountability in its deployment. Science stands at a threshold: the choices we make now will shape not only the future of knowledge production, but also the societal role and legitimacy of science itself.

EXISTING APPLICATIONS OF GENERATIVE AI IN SCIENCE

The integration of generative artificial intelligence into scientific research has rapidly evolved from a theoretical ambition to a transformative force that is reshaping methodologies across disciplines. Unlike conventional computational tools, generative AI systems — such as large language models (LLMs), deep generative networks, and reinforcement learning agents — are capable of producing novel content, generating hypotheses, and simulating complex systems. These capabilities position generative AI not just as an assistant in science, but as a potential co-creator of scientific knowledge.

One of the most celebrated applications of generative AI in recent years is the breakthrough in protein structure prediction achieved by DeepMind's AlphaFold (Jumper *et al.*, 2020, 2021). Using deep learning techniques, AlphaFold demonstrated the capacity to predict 3D protein structures with unprecedented accuracy, solving a decades-old challenge in molecular biology. This milestone has not only accelerated basic biological research but also opened new avenues for drug design and therapeutic innovation.

In pharmaceutical research more broadly, generative AI is being deployed to explore vast chemical search spaces efficiently. AI models are used to generate candidate molecules with desired properties, evaluate their potential through simulation, and even design synthetic pathways. For example, AI systems have identified new antibiotics (Awan *et al.*, 2024) and antimalarial compounds that were previously unknown (Lin *et al.*, 2024), reducing the time and cost traditionally associated with drug discovery pipelines.

In materials science, generative models assist researchers in identifying new compounds with specific mechanical, thermal, or electrical properties (Chen *et al.*, 2024; Pogue *et al.*, 2023). These models enable inverse design — starting from a set of desired properties and working backwards to discover materials that exhibit those characteristics. This approach contrasts with traditional trial-and-error methods and allows for the accelerated development of advanced materials in fields such as semiconductors, batteries, and nanotechnology.

Climate science, too, has seen significant advances through generative AI. Models are used to simulate weather patterns (Lam, 2023), downscale global climate projections to regional levels, and improve early-warning systems for extreme weather events (Reichstein *et al.*, 2025). AI-enhanced simulations offer faster and more precise results, allowing researchers and policymakers to better anticipate environmental risks and plan mitigation strategies. The coupling of generative AI with high-resolution earth observation data, such as in the EU's "Destination Earth" initiative, exemplifies the potential for AI to inform sustainable development.

Generative AI is also facilitating progress in the humanities and social sciences. In historical research and digital humanities, language models trained on archival corpora assist in analysing text patterns (Assael *et al.*, 2022), translating ancient languages (Gutherz *et al.*, 2023), and reconstructing lost documents (Weber, 2024). These tools make it possible to interrogate vast repositories of textual data and many other artifacts in ways that were previously infeasible. Meanwhile, in sociology and political science, generative AI helps model and simulate social behaviours (Park *et al.,* 2024), enabling scenario testing and the development of evidence-informed policy.

Beyond the actual generation process, generative AI also enhances the dissemination of scientific knowledge. AI systems are used to summarise research articles, translate findings into lay terms, and even assist in drafting publications. Tools such as natural language generation models contribute to the peer review process by detecting inconsistencies, assessing language clarity, or flagging possible plagiarism. While these applications offer efficiency, they also raise important questions regarding authorship, accountability, and the integrity of scientific communication. Many conferences are now prohibiting the use of AI tools in the review process, a rule that is in effect very hard to enforce and whose results are, in any case, mixed.

In many areas of application, generative AI is already reshaping scientific practice across domains. From automating routine tasks to co-generating novel hypotheses and experimental designs, its influence is far-reaching and growing. The full potential of these technologies can only be realised through responsible, inclusive, and transparent integration into scientific workflows. By embedding generative AI into the fabric of research while upholding scientific norms and values, we move toward a future in which machines not only serve science but help reimagine its very possibilities.

PREDICTABLE FUTURE DEVELOPMENTS

As artificial intelligence continues to mature, several technological trends and methodological breakthroughs are poised to reshape scientific research in the near term. These developments promise not only to enhance the efficiency and creativity of scientific workflows, but also to redefine the epistemic boundaries of what constitutes “doing science”. Among the most significant of these trends are the emergence of agentic AI systems, the increasing adoption of retrieval-augmented generation (RAG) and related techniques, the refinement of prompting strategies such as chain-of-thought (CoT), among many other new approaches to the design of AI models targeted for scientific development.

Agentic AI represents a new paradigm in the integration of AI with scientific reasoning and experimentation (Acharya *et al.*, 2025). Rather than acting as passive tools that respond to individual prompts, agentic systems are autonomous or semi-autonomous AI entities capable of decomposing complex tasks into subgoals, executing sequences of operations, and adapting their strategies based on feedback. These agents, often orchestrated through frameworks like AutoGPT (Fırat *et al.*, 2023), MemGPT (Packer *et al.*, 2023) or LangChain (Topsakal *et al.*, 2023), can perform literature searches, simulate experimental setups, and even propose novel hypotheses. These developments point toward semi-autonomous digital entities capable of complex tasks, a direction aligned with a vision of AI as a collaborative partner in research.

In tandem with these developments, retrieval-augmented generation (RAG) techniques are becoming foundational in scientific applications. RAG (Lewis *et al.*, 2020) enhances the reliability and factual grounding of generative models by combining their text generation capabilities with dynamic access to structured or unstructured knowledge bases of even to specific engines for particular domains, such as physics, mathematics or chemistry. The use of factual data coming from outside the model is especially important in scientific domains, where factual accuracy and traceability are paramount. By retrieving relevant context at inference time - such as datasets, scientific articles, models or experimental protocols - RAG mitigates the hallucination problem that plagues large language models and provides users with verifiable sources and citations. Both epistemic transparency and reproducibility are reinforced by RAG-based approaches, which already represent the standard today in scientific searches.

All these techniques led to the appearance of several "deep research" systems, powered by companies like Google, OpenAI and Perplexity. But perhaps one of the most compelling visions of near-term AI evolution comes from recent work by Sakana AI (Yamada *et al.*, 2025). Sakana's model departs from traditional approaches by using evolutionary computation principles — specifically, techniques inspired by biological diversity and ecological interactions — to continuously generate and select high-performing models. Rather than statically training a single monolithic model, their framework generates populations of models that adapt to different tasks or constraints. An article generated using this methodology was recently accepted in the workshop of a major conference in AI (ICRL 2025, the International Conference on Learning Representations) under a double blind review process, illustrating the viability of new scientific results fully generated by machines.

These predictable near-term developments point to a broader transformation of the scientific enterprise. AI will cease to be a background tool and will increasingly become a collaborative agent in the production of knowledge. But with this promise comes a responsibility. These systems must be developed and integrated in ways that preserve the transparency, reproducibility, and ethical integrity of science. This includes training researchers in the use of AI, ensuring equitable access to resources such as high-performance computing and high-quality datasets, and designing regulatory frameworks that foster innovation without compromising societal values.

ETHICAL ISSUES

As artificial intelligence becomes increasingly embedded in scientific practice, its ethical implications come into sharper focus. The application of AI in science holds vast potential for accelerating discovery, improving reproducibility, and enabling new forms of interdisciplinary collaboration. However, these benefits are accompanied by significant ethical challenges that must be addressed if AI is to serve the long-term interests of science and society.

One of the most pressing concerns is data integrity and bias. AI systems, particularly those based on machine learning, are fundamentally reliant on the quality of their training data. Incomplete, skewed, or non-representative datasets can lead to models that perpetuate existing scientific or societal biases. For instance, if biomedical models are trained predominantly on data from certain populations, they may produce less accurate results for underrepresented groups. This raises serious questions about fairness, scientific generalizability, and social responsibility that have to be tackled (Yang *et al.,* 2024). The SAPEA report underscores the importance of ensuring that datasets used in scientific AI applications are diverse, representative, and well-documented.

Closely related is the issue of reproducibility and transparency. Many AI systems, especially those built on deep learning architectures, are characterised by high performance but low interpretability. These so-called “black box” models pose a challenge to the scientific norm of transparency, where results must be explainable and open to scrutiny. If researchers cannot trace how an AI system reached a given conclusion, it becomes difficult to validate or reproduce the findings. The SAPEA report argues strongly for the adoption of interpretable AI models in scientific contexts or, at the very least, the development of standards and tools that make their decision-making processes more understandable. Both well established and very recent results can be used to improve interpretability (Bereska *et al.*, 2024; Lundberg *et al.*, 2017) but the issue remains an important one.

Another growing ethical concern is the generation of false or misleading content. Generative AI, while valuable for summarising data, drafting reports, or generating hypotheses, can also produce plausible-sounding but incorrect or fabricated results. This risk is particularly acute in low-supervision environments or in the hands of non-experts. The boundary between genuine scientific output and AI-generated fiction can become blurred, potentially undermining public trust in science. SAPEA recommends mechanisms to label AI-generated content clearly and to integrate verification layers into scientific workflows. Furthermore, irrelevant and automatically generated articles can rapidly overwhem the ability of the scientific community to identify relevant contributions by using peer review.

Academic integrity and authorship are also being tested by the rise of AI tools. While AI can support the writing and editing of scientific texts, it raises questions about attribution and intellectual ownership. Who is the author when an AI system generates significant portions of a manuscript? How should contributions from AI be acknowledged in peer-reviewed publications? There is a clear need for new conventions and guidelines to distinguish between human and machine-generated contributions, ensuring that credit and accountability are fairly assigned.

Finally, there is the risk of dual use and misuse of AI in scientific domains. AI models designed for beneficial purposes, such as protein folding prediction or genomic editing, could be repurposed for harmful applications, including bioweapons development, attack drones or surveillance. The scientific community must remain vigilant and proactive in anticipating these risks. Ethical review processes, responsible research and innovation frameworks, and international collaboration are essential to navigate these complex challenges.

CONCLUSIONS

The integration of artificial intelligence, and particularly of generative AI, into scientific practice represents a paradigm shift in how knowledge is produced, validated, and disseminated. From protein folding and drug discovery to climate modelling and historical text analysis, AI systems have demonstrated their capacity to accelerate innovation, enable interdisciplinary collaborations, and support the generation of new scientific hypotheses. As the European Union and the global research community explore these new capabilities, it is clear that AI is no longer merely a tool for computation as it is becoming a collaborator in the scientific enterprise.

However, this transformation is not without challenges. The opacity of many AI models, the potential for biased or non-reproducible outputs, and the risk of misuse all point to the necessity of a principled and cautious integration of AI into science. Ethical issues, anging from data integrity and authorship to environmental sustainability and global equity, must be addressed through robust governance frameworks, transparent methodologies, and inclusive infrastructures.

The SAPEA evidence review provides a roadmap for such a transition, calling for strategic investments in trustworthy AI, the establishment of new institutions such as EDIRAS and an AI Science Council, and the promotion of training and resources to ensure broad and equitable participation. Predictable near-term developments, such as agentic AI systems, advanced RAG techniques, deep research, CoT prompting, and evolutionary model design, underscore the urgency of adapting both scientific practice and policy frameworks to a rapidly evolving technological landscape.

The successful and responsible adoption of AI in science requires a dual commitment: to technical excellence and to the enduring values of transparency, integrity, inclusiveness, and public trust. The choices made today will determine whether AI serves as a tool for accelerating discovery and solving global challenges, or as a force that deepens inequalities and erodes confidence in scientific knowledge. Science and its practioners must rise to this challenge, not only with innovation but with wisdom.

Notes

The author was one of the experts involved in the drafting of the SAPEA evidence report, which provided a significant basis for this article. Artificial intelligent tools, including GPT4o, Gemini, Perplexity, Google and NotebookLM were used in the production of this article, namely in the research, summarization, and drafting phases.

COMUNICAÇÃO APRESENTADA ÀS CLASSES DE CIÊNCIAS E DE LETRAS
NA SESSÃO CONJUNTA DE 9 DE JANEIRO DE 2025

COMUNICAÇÃO RECEBIDA A 6 DE JUNHO DE 2025

REFERENCES

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

Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & De Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature*, *603*(7900), 280-283.

Awan, R. E., Zainab, S., Yousuf, F. J., & Mughal, S. (2024). AI-driven drug discovery: Exploring Abaucin as a promising treatment against multidrug-resistant *Acinetobacter baumannii*. *Health Science Reports*, *7*(6), e2150.

Bereska, L., & Gavves, E. (2024). Mechanistic Interpretability for AI Safety--A Review. *ArXiv Preprint ArXiv:2404.14082*.

Chen, C., Nguyen, D. T., Lee, S. J., Baker, N. A., Karakoti, A. S., Lauw, L., Owen, C., Mueller, K. T., Bilodeau, B. A., & Murugesan, V. (2024). Accelerating computational materials discovery with machine learning and cloud high-performance computing: from large-scale screening to experimental validation. *Journal of the American Chemical Society*, *146*(29), 20009-20018.

Fırat, M., & Kuleli, S. (2023). What if GPT4 became autonomous: The Auto-GPT project and use cases. *Journal of Emerging Computer Technologies*, *3*(1), 1-6.

Gutherz, G., Gordin, S., Sáenz, L., Levy, O., & Berant, J. (2023). Translating Akkadian to English with neural machine translation. *PNAS Nexus*, *2*(5), pgad096.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., & Potapenko, A. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583-589.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Tunyasuvunakool, K., Ronneberger, O., Bates, R., Žídek, A., & Bridgland, A. (2020). AlphaFold 2. *Fourteenth Critical Assessment of Techniques for Protein Structure Prediction*, 13.

Lam, R. (2023). GraphCast: AI Model for Faster and More Accurate Global Weather Forecasting. *Google DeepMind Blog*.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., & Rocktäschel, T. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, *33*, 9459-9474.

Lin, M., Cai, J., Wei, Y., Peng, X., Luo, Q., Li, B., Chen, Y., & Wang, L. (2024). MalariaFlow: A comprehensive deep learning platform for multistage phenotypic antimalarial drug discovery. *European Journal of Medicinal Chemistry*, *277*, 116776.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Packer, C., Fang, V., Patil, S., Lin, K., Wooders, S., & Gonzalez, J. (2023). *MemGPT: Towards LLMs as Operating Systems*. https://arxiv.org/pdf/2310.08560

Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., Morris, M. R., Willer, R., Liang, P., & Bernstein,

M. S. (2024). Generative agent simulations of 1,000 people. *ArXiv Preprint ArXiv:2411.10109*.

Pogue, E. A., New, A., McElroy, K., Le, N. Q., Pekala, M. J., McCue, I., Gienger, E., Domenico, J., Hedrick, E., & McQueen, T. M. (2023). Closed-loop superconducting materials discovery. *NPJ Computational Materials*, *9*(1), 181.

Reichstein, M., Benson, V., Blunk, J., Camps-Valls, G., Creutzig, F., Fearnley, C. J., Han, B., Kornhuber, K., Rahaman, N., & Schölkopf, B. (2025). Early warning of complex climate risk with integrated artificial intelligence. *Nature Communications*, *16*(1), 2564.

SAPEA. (2024). *Successful and timely uptake of artificial intelligence in science in the EU: evidence review report* . SAPEA. doi: 10.5281/zenodo.10849580

Topsakal, O., & Akinci, T. C. (2023). Creating large language model applications utilizing Langchain: A primer on developing LLM apps fast. *International Conference on Applied Engineering and Natural Sciences*, *1*(1), 1050-1056.

Weber, T. (2024). *Inside the AI Competition That Decoded an Ancient Herculaneum Scroll | Scientific American*.

Yamada, Y., Lange, R. T., Lu, C., Hu, S., Lu, C., Foerster, J., Clune, J., & Ha, D. (2025). The AI scientist-v2: Workshop-level automated scientific discovery via agentic tree search. *ArXiv Preprint ArXiv:2504.08066*.

Yang, Y., Lin, M., Zhao, H., Peng, Y., Huang, F., & Lu, Z. (2024). A survey of recent methods for addressing AI fairness and bias in biomedicine. *Journal of Biomedical Informatics*, 104646.