Scientific Advice Mechanism
to the European Commission

April 2024

SAPEA evidence review report

Successful and timely uptake of artificial intelligence in science in the EU

About SAPEA

SAPEA's role in the Scientific Advice Mechanism is to provide independent, high-quality reviews of the evidence to inform the policy recommendations made by the Group of Chief Scientific Advisors.

We are a consortium of Academy Networks, funded by Horizon Europe, representing a large number of academies from different countries.

Through these Networks, we bring together outstanding expertise from natural sciences, engineering and technology, medical, health, agricultural and social sciences, and the humanities.

About this report

In July 2023, the College of Commissioners asked the Scientific Advice Mechanism of the European Commission to provide evidence-based advice on how to accelerate the responsible uptake of AI in science.

To address this question, SAPEA assembled an independent, international, and interdisciplinary working group of leading experts in the field, nominated by and selected from European academies and their Networks. Between October 2023 and January 2024, the working group reviewed and compiled the latest evidence on the subject to create an evidence review report. This report informs the accompanying Scientific Opinion of the Group of Chief Scientific Advisors, which contains the requested policy recommendations.

This brochure contains the executive summary of the evidence review report, summarising its main conclusions. To read the full report and accompanying Scientific Opinion, visit https://scientificadvice.eu/advice/artificial-intelligence-in-science/.

This SAPEA evidence review report gathers the relevant scientific evidence to analyse:

> How can the European Commission accelerate a responsible uptake of AI in science (including providing access to high-quality AI, respecting European Values) in order to boost the EU's innovation and prosperity, strengthen the EU's position in science, and ultimately contribute to solving Europe's societal challenges?

Specifically, the report approaches the topic through the lens of AI's impact on:

- scientific process, including the underlying principles upon which the scientific endeavour is organised
- people, including the skills, competencies, and infrastructure needed by scientists of tomorrow
- policy design, in the context of ensuring a timely, responsible, and innovative uptake of AI in science in Europe

In the rapidly-evolving field of AI, there is no universally accepted definition, nor a clear taxonomy outlining its various branches. Establishing such a definition would facilitate international collaboration among different countries. Therefore, recently, OECD countries have agreed to define an AI system as:

> a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

As AI applications permeate across many sectors, including in research, it is imperative that the EU takes hold of the opportunities, acts upon the challenges, and safeguards citizens from the risks that this fast-evolving technology can bring. As a companion effort to the EU's regulatory AI Act in progress, which is working

to promote the uptake of human-centric and trustworthy AI while ensuring a high level of protection of health, safety and fundamental rights (per AI Act Recital 1) in Europe, the European Commission aims to understand the specifics of AI technology not only developed by science, but as applied to science; that is, AI in science. This report reviews current evidence and potential policies that could support the responsible and timely uptake of AI in science in the EU that may enhance the EU's innovation and prosperity. The scope of this report is confined to the takeup of AI in scientific research, rather than its implications for society more generally. In particular, it does not address the manifold and very significant challenges that arise from the growing and rapid deployment of AI technologies in specific social domains, and which fall outside the scope of this report and which the EU's AI Act is intended to address.

Landscape of AI in research and innovation

Computational power

The computational power ('compute') required for advanced machine learning systems has increased exponentially for many decades and in particular since 2010, ultimately leading to a divide between academia and industry regarding access to specialised software, hardware, and skilled workforce. Academic institutions released the most significant machine learning systems until 2014, but industry has now taken the lead. Generative AI models (especially large language models (LLMs) and diffusion models), convolutional neural networks for vision applications and models trained using deep reinforcement learning, widen the gap between industry and research due to the huge computational resources needed to train them. Governments are investing in computing capacity, but lag behind private

sector efforts. Newcomers, startups and AI research laboratories frequently build on big tech's cloud services to train models and launch products.

Data

Besides compute, data is a crucial resource for AI development. However, the need to comply with copyright laws and research ethics requirements create challenges for public institutions in obtaining and processing data. While efforts are being made to provide fair and equal access while preserving privacy and ownership rights, issues remain unresolved.

Geopolitics of AI

The USA, China and the UK (ranking third, but lagging far behind the first two) dominate AI research globally, but other European countries contribute significantly. China leads in scientific domains where AI plays a prominent role, while the USA excels in health-related fields and the EU in social sciences and humanities. In 2022, the USA had more authors contributing to significant machine learning systems than other countries like China and the UK. The top-cited AI papers are from companies like Google, Meta, and Microsoft. The growth rate of AI-related publications is much faster than overall scientific production globally.

AI investment is surging globally, primarily driven by private sector investment, with the USA taking the lead. The EU has lost innovation leadership due to low research and development (R&D) investment and fewer startups. The problem is the commercialisation of R&D and scaling up. European efforts to boost AI development include programmes such as Horizon 2020, Horizon Europe, the Large AI Grand Challenge within the AI innovation package, and access to supercomputing resources through the EuroHPC JU network: https://eurohpc-ju.europa.eu/.

Regulatory landscape

Numerous AI-specific legal and regulatory measures are emerging worldwide, with many nations yet to enact comprehensive AI legislation. Most countries rely on existing frameworks for regulation, complemented by governance guidelines. As of October 2023, 31 countries had enacted AI laws, while an additional 13 countries are discussing potential regulations. The EU and China are leading in developing comprehensive AI regulations. China has been at the forefront of AI regulation, enacting specific measures for algorithmic bias, the responsible use of generative AI, and more robust oversight of deep synthesis technology. In the EU, considerable attention has been devoted to its AI Act, described by the European Commission as the most comprehensive AI legislation in the world. More recently, USA AI policy has seen the publication of several legally-mandated reforms.

Opportunities and benefits of AI in science

The increasing accessibility of generative AI and other machine learning tools for the analysis of large volumes of data has led scientists across various disciplines to incorporate them into their research. These tools facilitate the analysis of large amounts of text, code, images, and field-specific data, enabling scientists to generate new ideas, knowledge, and solutions. The number of scientific projects incorporating AI proliferates, with successful examples in protein engineering, medical diagnostics, and weather forecasting. Beyond facilitating groundbreaking discoveries, AI is also transforming the daily academic work of scientists, from supporting manuscript writing to code generation.

Accelerating discovery and innovation

AI's transformative potential extends to accelerating scientific discovery and innovation. The vast amount of research knowledge in natural language format is harnessed through literature-based discovery processes, using existing literature in scientific papers, books, articles, and databases to produce new knowledge. Researchers can now use LLMs to mine scientific publication archives to generate new hypotheses, develop research disciplines, and contextualise literature-based discovery. They can also use advanced search methods, such as those based on deep reinforcement learning, to comb vast search spaces, opening the way to AI-driven discoveries.

Scientific domains relying on large amounts of data seem to have taken up AI to a larger extent in their research processes. The generation of Big Data in these research fields presents a challenge that AI is well-equipped to address. AI algorithms analyse massive, complex, and high-dimensional datasets, enabling researchers to identify patterns and develop new insights. In fields like astronomy, particle physics, and quantum physics, where even a single experiment generates vast amounts of data, AI algorithms identify patterns at scale with increased speed, allowing scientists to discover never-before-seen patterns and irregularities. AI is becoming an indispensable tool for extracting knowledge from experimental data.

AI and machine learning tools can also help bridge the gap between diverse research fields, promoting cross-disciplinary collaborations. By incorporating Big Data analytics using AI, humanities researchers incorporate quantitative measures, diversifying their research and research questions. For example, some historians use machine learning tools to examine historical documents by analysing early prints, handwritten documents, ancient languages, and dialects.

Furthermore, AI exhibits potential in advanced experimental control of large-scale complex experiments. For example, physicists are now incorporating AI systems that use reinforcement learning to gain better control over their experiments.

Automating workflows

Traditionally, researchers performed experiments manually, often involving labour-intensive tasks. However, technological advancements enable the automation of a significant portion of experimental workflows. AI is revolutionising experimental simulation and automation, opening up new possibilities for research.

Enhancing output dissemination

AI is also enhancing the dissemination of research outputs. AI-powered language editing empowers non-native English speakers to refine their scientific manuscripts, bridging communication gaps between experts and the public. AI can also simplify the publishing process for newcomers, potentially fostering more inclusive scientific discourse.

Challenges and risks of AI in science

In taking up AI, scientific researchers need to address bias, respect principles of research ethics and integrity, and deal responsibly with issues surrounding reproducibility, transparency, and interpretability.

Limited reproducibility, interpretability, and transparency

The use of AI in science compounds existing concerns about reproducibility, while the opacity of AI algorithms also poses significant challenges to scientific integrity, interpretability and public trust. AI algorithms can generate useful outputs, but their opaque nature makes it difficult to verify the accuracy and validity of

research findings. The lack of transparency in AI algorithms hinders reproducibility, as researchers cannot replicate important discoveries without knowledge or understanding the underlying methodological processes. The opacity of AI algorithms raises concerns about accountability and trust, particularly in high-stakes applications such as healthcare.

The increasing prevalence of generative AI models and computer vision systems produced by industry raises concerns about their opacity and the lack of control over human evaluation by academic researchers. Insufficient access to scalable pipelines, large-scale human feedback, or data hinders academic researchers' ability to assess the safety, ethics, and social biases of machine learning models. The challenge of building state-of-the-art AI models due to the scarcity of computational and engineering resources leads to a reliance on commercial models, limiting reproducibility and advancement outside of commercial environments. The monopolisation of AI capabilities by tech giants raises concerns about their control over the development and application of AI, potentially limiting scientific progress and ethical considerations.

Poor performance (inaccuracy)

Despite their remarkable capabilities, AI models are susceptible to performance issues arising from various factors. One such factor is the quality of training data. The model's predictions will inevitably suffer if the data used to train an AI model is biased, inaccurate, or incomplete. Additionally, AI models require ongoing updates to maintain their accuracy. Failure to retrain models with current data can lead to outdated algorithms generating inaccurate outcomes.

Another crucial aspect is the representativeness of training data. AI models often learn from data that does not accurately reflect real-world populations. This discrepancy can introduce biases into the model, resulting in erroneous predictions.

Finally, the lack of adequate knowledge and training among researchers and developers contributes to performance shortcomings. AI models may be developed and deployed irresponsibly without proper expertise, leading to ethical and legal complications and concerns.

Fundamental rights protection and ethical concerns also arise. AI has the potential to perpetuate existing social biases and discrimination because AI systems trained on historically biased data and thus likely to reproduce these biases in their outputs. This can have a negative impact on people from marginalised groups, who may be unfairly discriminated against. AI systems can also introduce new forms of bias, such as visual perception bias. Machine vision systems may be biased because they are trained on datasets not representative of the real world.

In AI research, industry is now racing ahead of academia. Industry research has greater access to resources, such as data, talent, computing power, infrastructure, and funding, enabling them to take the lead over academia in developing sophisticated AI systems. This can disadvantage smaller institutions and academic researchers, making it more difficult for them to advance research.

AI systems can also raise privacy and data protection concerns since they often collect and process personal data and other, confidential information. There are several other challenges to advancing AI in science e.g. its adverse environmental impacts.

Misuse (malicious actors) and unintended harm

The misuse of AI in scholarly communication can lead to several significant social harms, including the proliferation of misinformation, the creation of low-quality outputs, and plagiarism. It constitutes research misconduct.

AI-generated content can be challenging to distinguish from human-generated content, increasing the risk of spreading misinformation. Predatory journals and paper mills can use AI to create fraudulent research papers. AI can make it easier to plagiarise content, potentially violating copyright and other intellectual property rights. The ease of producing AI-generated content may lead to an increase in the number of irrelevant papers. This can erode trust in scientific findings. AI-based tools can falsify information, which could lead to research misconduct. Using AI to generate content may lower the bar on the required scientific quality of the original work.

Using AI-based tools can automate specific tasks in the peer review process but, unlike human reviewers, cannot properly assess the novelty and validity of research findings reviewers can. As of today, AI still performs poorly in attempting to assess research quality, lacking human reviewers' deep knowledge, capability of grasping meaning, significance and human understanding. Using AI-based systems to evaluate scientific research may introduce bias and additional errors into the research assessment process.

Societal concerns

The advancement of AI has raised concerns about its potential impact on society. One concern is the unfair appropriation of scientific knowledge, as large tech companies increasingly leverage scientific talent from universities and volunteer developers' contributions from public code hosting and community platforms. Simultaneously, these companies hold patents and profits for themselves, while controlling access to computing and datasets. Additionally, using copyrighted material as training data for AI models raises concerns about copyright infringement, yet those whose IP rights may have been interfered with lack the capacity or resources to challenge purported infringement and seek redress, while

identifying how IP law should apply in these contexts remains unsettled and uncertain.

Another concern is AI's potential to manipulate and spread misinformation at scale. Additionally, it may pose cybersecurity threats, including malware generation through unsafe code with bugs and vulnerabilities, advanced phishing attacks using LLMs for large-scale deployment, cybercriminals leveraging AI tools for malicious activities or deepfakes and voice cloning leading to impersonation, fraudulent digital content generation and realistic voice scams. Furthermore, AI may impact modern warfare and facilitate bioweapons development.

Impact on scientists and researchers

Research environments, literacy and training

AI can change the research context and environment, automating tasks, enhancing productivity, and liberating researchers from menial tasks. It can also amplify a researcher's expertise by personalising research tools and tailoring support and assistance to individual needs, preferences, and expertise. This transformation demands adaptation and the acquisition of new skills. To benefit fully from AI, universities and researchers must invest in AI literacy and digital skills, foster a collaborative culture between humans and AI in the framework of human-centred AI, and embrace the dynamic interplay between human expertise and AI augmentation.

AI literacy involves understanding the concepts, abilities, and limitations of AI technology and being able to effectively communicate with it while evaluating its trustworthiness. Ethical awareness, critical thinking, value addition to AI output, and fact-checking are other crucial skills for successful AI integration in research.

Adapting to the rapidly changing research environment is crucial for remaining competitive in the field. Several AI teaching programmes exist in Europe to address these needs. These aim to educate individuals in various aspects of AI, from technical knowledge to ethical considerations to help develop a skilled workforce capable of addressing the growing demand for AI expertise across industries and sectors, including research.

Inequalities and biases

AI's potential to transform research demands a conscious effort to address the geographical disparities in AI access and development and gender imbalances. Researchers should embrace a human-centred approach, mitigate biases, and collaborate with stakeholders to ensure AI's positive and equitable impact on society.

Impact on researchers

Adopting AI in research careers may lead to negative consequences, undermining mental well-being, increasing job insecurity, pressure, and unfair discrimination. Additionally, using AI for review and selection processes can erode a sense of belonging among researchers. It is important to address these challenges to ensure the appropriate implementation of AI in research.

Evidence-based policy options

Based on these findings, this report identifies five broad challenges that confront EU policymakers that may help to accelerate the responsible and timely uptake of AI in scientific and research communities, thereby supporting European innovation and prosperity. In this context, 'responsible' is taken to mean that accelerated uptake of AI should strive to be in accordance with the foundational commitments of

scientific research and the foundational values underpinning the EU as a democratic political community and thus ruled by law, ensuring respect for the fundamental rights of individuals and the principles of sustainable development.

The primary challenge that must be addressed in order to accelerate the uptake of scientific research both in AI, and using AI for research, concerns resource inequality between public and private sector research in AI. To foster scientific uptake of AI responsibly, four further challenges must be addressed, concerning:

- scientific validity and epistemic integrity
- opacity
- bias, respect for legal and fundamental rights and other ethical concerns
- threats to safety, security, sustainability, and democracy

This report then sets out a suite of policy options which are directed towards addressing one or more of these challenges. These policy proposals include:

- founding a publicly funded EU state-of-the art facility for academic research in AI, while making these facilities available to scientists seeking to use AI for scientific research, thereby helping to accelerate scientific research and innovation within academia
- fostering research and the development of best practices, benchmarks, and guidelines for the use of AI in scientific research aimed at ensuring epistemic integrity, validity and open publication in accordance with law and conducted in an ethically appropriate manner
- developing education, training, and skills development for researchers, supplemented by the creation of attractive career options for early career AI researchers to facilitate retention and recruitment of talented AI researchers within public research institutions

- developing publicly-funded, transparent guidelines and metrics, using them as the basis for independent evaluation and ranking of scientific journals by reference to their adherence to principles of scientific rigour and integrity. The publication of these evaluations and rankings would be intended to provide a more thorough, rigorous, informed, and transparent indication of the relative ranking of scientific journals in terms of their scientific rigour and integrity than existing market-based metrics devised by industry, helping to identify predatory and fraudulent journals
- establishing an EU 'AI for social protection' institute, which engages in information exchange and collaborates with other similar public institutes concerned with monitoring and addressing societal and systemic threats posed by AI in Europe and globally, proactively monitoring and providing periodic reports and making recommendations aimed at addressing threats to safety, security, sustainability, and democracy

To read the full report and accompanying Scientific Opinion, visit https://scientificadvice.eu/advice/artificial-intelligence-in-science/.

scientificadvice.eu
@EUScienceAdvice

Contact us
EC-SAM@ec.europa.eu

Within the Scientific Advice Mechanism, SAPEA is funded by the European Union.
The activities of associated partners Academia Europaea and Cardiff University
are funded by UKRI (grant number 10033786).