Scientific Advice Mechanism
to the European Commission

26–27 October 2023

Foresight workshop report

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

- SAPEA (2024). *Successful and timely uptake of artificial intelligence in science in the EU: Foresight workshop report*. Berlin: SAPEA.
- DOI 10.5281/zenodo.10974075
- Downloadable from https://scientificadvice.eu/advice/artificial-intelligence-in-science/

Version history

Version	Date	Summary of changes
1.0	15 April 2024	First published version

Scientific Advice Mechanism
to the European Commission

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

Foresight workshop report

26–27 October 2023

Table of contents

About the Scientific Advice Mechanism

The Scientific Advice Mechanism provides independent scientific evidence and policy recommendations to the European institutions by request of the College of Commissioners.

We consist of three parts: the Group of Chief Scientific Advisors, SAPEA, and the SAM secretariat.

About the Advisors

The Group of Chief Scientific Advisors consists of seven eminent scientists whose role in the SAM is to make policy recommendations. They are chosen for their outstanding level of expertise, covering a wide range of scientific fields.

About SAPEA

SAPEA is a consortium of European academy networks representing around 110 academies from across Europe. Its role in the SAM is to provide independent, high-quality reviews of the evidence to inform the policy recommendations made by the Group of Chief Scientific Advisors.

About the SAM secretariat

The SAM is supported by the SAM secretariat, a small team within the European Commission which acts as a liaison between the other parts of the SAM and the European Commission, including the Commissioner for Research, Science and Innovation.

About the report

In July 2023, Margrethe Vestager, Executive Vice-President of the European Commission and acting Commissioner for Innovation, Research, Culture, Education and Youth, asked the Group of Chief Scientific Advisors to deliver advice on the topic of successful and timely uptake of artificial intelligence in science in the EU. Key area 1 of the scoping paper, on vision and foresight, was requested to be addressed by the end of 2023.[1]

To address this key area, the Scientific Advice Mechanism organised a foresight workshop. The workshop brought together an interdisciplinary group of 27 scientists, experts, technologists and innovative thinkers to explore what the future of AI-enabled science could look like in 2035, in answer to the main question:

> What impetus could AI give to scientific productivity/innovation and what benefits, incentives and challenges would AI-enabled research bring to the European innovation ecosystem and society as a whole?

This workshop took place on 26 — 27 October in person in Brussels. It was organised by the SAM, in partnership with the School of International Futures (SOIF). Expert participants generated ideas, debated key issues and imagined possibilities for the future. This report summarises the exchanges, main ideas and discussions. The description and rationale for the exploratory foresight approach and the names of expert participants can be found in the appendices.

Disclaimer

This report summarises the discussions, ideas and points of view of 27 participants based on their expertise on and experience with AI. This report is not a review of evidence, and it is not part of the evidence review process of the SAPEA working group on the topic. The report is written in a non-attributed style and does not represent the opinions of individual participants involved.

Introduction

AI gained instant prominence in mainstream media with ChatGPT, which broke the record for the fastest-growing app in history in January 2023.[2] However, early implementations of AI go back to the mid-1950s and the Logic Theorist (Newell & Simon 1956).[3] As for ChatGPT, science-fiction writer

1 Key areas 2, 3 and 4 of the Scoping Paper were addressed separately by SAPEA, through a process of evidence review supported by a working group.

2 *ChatGPT sets record for fastest-growing user base - analyst note*, Krystal Hu, Reuters (2023). https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

3 A. Newell and H. Simon, *The logic theory machine—A complex information processing system*. In IRE Transactions on Information Theory, vol. 2, no. 3, pp. 61-79, September 1956, doi: 10.1109/TIT.1956.1056797.

Ted Chiang called it "a blurry JPEG of the web".[4] As such, both ChatGPT and other modern AI models reproduce the biases and power relations embedded in their corpora. In real life, too, a handful of the loudest voices dominate the discussion around AI, diminishing and stifling contributions from others.

In a foresight exercise, experts and leaders in AI from across Europe gathered to imagine and describe possible futures for AI-enabled research in the European research and innovation ecosystem, as well as to reflect on opportunities and challenges for reaching a desired future.

In this report, we gather the experts' ideas and reflections in an attempt to answer the following key questions:

- Are we on a sustainable growth course of AI that will ensure the EU's global competitiveness, safeguarding it from dependence on the technology from foreign innovation?
- Are we ensuring a responsible underlying framework of the AI that will imbue the European values in our own implementations?

As AI applications in science are growing exponentially, this is an opportune time to ensure that AI-enabled research in the EU is respecting the safety and autonomy of its citizens and its natural environment, while maximising the benefits of a technologically savvy and literate continent.

The AI impetus on scientific practices and productivity

Scientific productivity through scientific curiosity

Participants explored the future of AI-enabled research by the year 2035. They suggest that by this time, AI is unlikely to be creative, meaning it won't come up with genuinely new ideas such as new mathematical theorems. However, since AI output is derivative, it can potentially be used to test ideas and hypotheses or analyse data much faster than ever before. By 2035, AI could become a widespread tool across sciences, used to support scientists in reframing theories, thinking out of the box, generating ideas that are not obvious reiterations of the previous ones. Curiosity and creativity would remain singularly human faculties in the domain of scientists — at least within the defined time frame. AI may never reach 'human-level curiosity' understood as out-of-the-box thinking, and algorithmic reasoning may be fundamentally at odds with AI. Therefore, AI and human intellect can either strengthen each other or compete with each other.

Some participants also suggested that the capacity of AI to assist in writing, literature review, and peer review might greatly impact scientific processes. AI may even offer personalised daily literature summaries. AI could assist PhD supervision by offering personalised mentorship to students. AI-

4 *ChatGPT Is a Blurry JPEG of the Web*, Ted Chiang, The New Yorker (2003). https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web

enabled science could shift from a focus on publications to a focus on datasets. Data may supplant publications as the major type of scientific contributions. The 'publish or perish' career model may vanish. Unburdened by menial tasks, scientists could return to a more general practice of science, such as through the exploration of ideas and writing.

Furthering this thought, participants speculated that in a possible future, fields of science that rely heavily on data — such as science, technology, engineering and mathematics subjects — might become co-opted by AI and consequently devalued in favour of less data-driven sciences such as the humanities. However, participants noted that the humanities currently play a minor role in the development, governance and legislation of AI.

An opportunity to embed European values in AI

Participants identified the rapid development and implementation of AI in the EU as an opportunity to ensure European values in the development and deployment of AI systems. This includes working towards robustness and fairness, in addition to efficiency. Some mechanisms to achieve these goals could be based on existing EU policies such as the General Data Protection Regulation (GDPR) and Open Access.

Similarly to the achievements of the EU in setting strong data privacy standards, GDPR could be viewed as model legislation. Participants suggested that this could be applied to AI as well. The EU could take the lead as a frontrunner in the defence of privacy, providing incentives, involving social scientists and decision makers in the process, and leveraging industry partnerships. Through regulation similar to open access**,** a mandated amount of budget for each project could be allocated for ensuring AI remains ethical and sustainable.

Ethical AI was viewed by participants as a challenge in itself. Indeed, the way AI models are trained today tends to calcify the ethics and biases embedded in their training datasets. In the future, participants suggested consolidating multidisciplinary efforts to find ways to update the ethical norms in AI as they inevitably change and evolve over time. Participants also suggested a potential role for the EU in championing the equivalent of fair trade for data training, regulating how the data is trained, by whom, and under what conditions. This would ensure that the data is representative, ethical, and respectful of the rights and interests of the data providers.

AI for solving the most complex challenges

In line with European values, participants suggested that European AI needs to be designed for humanity rather than efficiency, and collaboration instead of competition. AI excels at understanding patterns, which allows scientists to use it for insight into complex global challenges such as climate change. For example, AI could be used to help make agriculture more resilient and sustainable, and not just more efficient.

Trained on historical data, AI necessarily lags behind contemporary ways of thinking, reflecting the biases of yesterday. Participants proposed that the biases encoded in AI could be used as a window into societal biases, including those of scientists. Since AI reflects the patterns of our collective subconscious, social scientists could use AI models as a tool for investigating and identifying our biases. They should do so with caution, however, as the models might still surprise us by distorting the biases in ways that have little to do with human sensibilities.

The duality of scientists and AI in 2035

Many uncertainties underlie the future of AI-enabled research. In principle, the AI machine learning systems work with solid and established mathematical optimisation methods that can be open or closed to public scrutiny. But if they are closed, one cannot avoid manipulation in the use of these methods by the service providers. On top of that, such systems cannot give an explanation for the output that they produce. The fact that AI output cannot always be explained (the inherent characteristic of machine learning) will divide scientists into two camps: some will seek explanations on their own; others, similar to engineers, will test the AI-generated solutions if they work as intended and, if confirmed, adopt them.

The AI adopters would pioneer a new paradigm in science, marking a departure from theory-based toward discovery-based science. This shift could initially lead to a boom in scientific discovery, which would then make the scientist — in the sense we understand this word today — disappear. AI-adopter scientists may send their AI avatars to conference talks they are not able to attend and receive a summarised report of the event later. In the most extreme of the possible futures, some participants speculated that only a fraction of scientists would remain active as they are today, while others may pursue science as a hobby with the aid of AI (see "Story Box: A Scientist of Tomorrow" on page 21).

Challenges and incentives of AI-enabled research

Promoting transdisciplinary thinking

Participants noted that the average output of AI models might sink to the lowest common denominator, which may, in the future, encourage a uniform way of thinking and reduce the diversity brought forth by human curiosity and intellect. In the preferred future, AI would increase the capacity for lateral thinking, not shrink it.

One way to promote diverse thinking is through AI development accompanied and informed by deep human thought and a humanities-based approach, so it does not become stilted and narrow. Participants propose a strong role for the EU in incentivising transdisciplinary initiatives and

collaborations. While transdisciplinary work is more common in the private sector, academia has little incentive toward it because of its siloed culture and overemphasis on key performance indicators.

Adapted funding schemes

Participants in the workshop identified several opportunities for adapting the current funding scheme to better support the responsible uptake of AI in science in the EU. These inputs are summarised below.

Participants highlighted that current funding opportunities in academia are predominantly short-term, which discourages the exploratory research that is necessary for innovation. In the private sector, large companies invest in long-term research goals. Participants suggest following in their footsteps for AI-enabled scientific project funding, especially in the case of interdisciplinary and ambitious projects that could potentially solve major societal challenges.

Exploratory risk-taking initiatives have the potential to lead to scientific breakthroughs and innovation. Participants suggested further encouraging risk-taking in science, for example, by allocating a budget for 'crazy ideas' — high risk/high reward experiments that yield immense results if successful.

Participants also brought up the need to increase talent attraction and retention in the EU. A possible way forward would be to create a robust research infrastructure with enough professional and financial incentives to entice talents from across domains who would then use it for transdisciplinary exploration of ideas. (For more information, see "A desired future: The BRAIN Institute" on page 13.)

Finally, participants also raised the point of public-private partnership in AI innovation. Public and private sector scientists may collaborate on research projects based on open source data. Opportunities and incentives could encourage these collaborations. These incentives could include prioritising the creation of an open source digital infrastructure, which could convince private companies to get involved, ultimately benefiting both public and private science and innovation.

Acknowledging AI generated content

In line with ongoing discussions in the scientific community, participants felt there was a need for clear guidelines for acknowledging the use of AI tools in writing a scientific publication. They proposed, for example, taking a cue from research publications from natural sciences, a rule to designate a main author responsible for the publication.

At the same time, participants felt additional knowledge and resources are needed to create and deploy better detection mechanisms to identify AI use in research. These could curb potential AI misuse by predatory journals as well as incentivise quality over quantity in research. In the end, some futures also proposed that humans might become redundant in the publication process, provided that AI output may become as good as human-generated content.

Safeguarding data quality

Human-generated data are essential for AI training. And it cannot be just any data. The saying 'you are what you eat' holds true for AI as well, which is why participants highlighted that it would be essential to make sure that the sources for AI datasets are transparent and traceable. Only that could enable AI to expedite and improve the scientific process in a reliable manner.

When imagining the future of AI in science, participants also raised the important point of synthetic data. Ten years from now, human-generated data might become mixed with synthetic data, generated by AI. This may in turn raise doubts over the quality of the scientific corpus, and the quality of the AI, due to a lower-quality corpus. It might be difficult to tell the difference between research results produced by AI and those produced by humans without labelling each accordingly. Paradoxically, setbacks in AI research might be the result of lesser quality human research. Accumulation of low-quality data may eventually lead to lower-quality AI models.

In upcoming regulatory frameworks, participants suggest a provision to trace and label data with respect to its origin. Participants illustrated this point with an analogy to anti-doping regulation in sports. Such regulation could define permissible and impermissible uses of AI in various contexts, including in science.

Certifying input

The robustness of the data used in training AI models guarantees the quality of the output. Allowing unverified datasets into the models could lead to undesirable and potentially irrelevant results. To ensure robust input datasets, participants suggested establishing a European certification system for AI training datasets.

Participants also suggested that — in addition to being robust — training datasets should also conform to the FAIR principles of Findability, Accessibility, Interoperability, and Reusability.[5] At the moment, datasets do not have to be model-agnostic. This means that companies can vendor-lock them to specific formats, making them unusable in other contexts and environments.

Furthermore, participants also proposed that AI research models should be built for specific rather than general purposes. General-purpose models prove ineffective. Scientists and engineers require built-for-purpose models to make their work in fact more efficient.

Focusing on AI literacy and education

Participants acknowledged that ignorance begets fear, and AI is no exception. The field moves so fast that even experts struggle to keep up with it. For this reason, they discussed AI literacy and capacity building with scientists and citizen representatives at the helm as a crucial development in the future

5 Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. *The FAIR Guiding Principles for scientific data management and stewardship*. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

of AI-enabled science. Furthermore, they also reflected on the advantages of educational programmes to also target decision-makers, who need AI literacy and expertise to inform their regulations (see "Story box: How much can we rely on AI?" on page 20).

Participants reflected that the EU's traditionally siloed educational system would most likely change in the future. To support AI literacy and education, they proposed potential actions such as incentivising mainstream comprehensive education programmes for the public with the basics of STEM and statistics, which would raise awareness of AI, enhance its uptake in Europe, and ease integration in business.

Participants also highlighted that resources would be required to support such goals. AI education programmes would need to ensure access to relevant computational resources. In addition, policymakers and educators should make the technology broadly available outside the teaching context. The combination of education and propagation may diminish fears of AI and empower people to use it with care and understanding. Educational programmes can ask to tackle real-life global challenges, such as the topics of the EU's Missions programme. It would serve two purposes, both to educate and to provide a brainstorm for solutions.

Additionally, participants recognised that AI might be hacked and manipulated by malicious forces, which could lead to misinformation injection and contaminate AI output without anybody knowing. For this reason, AI literacy would be crucial, in addition to AI being regulated by law, and AI output being verified and overseen by humans.

A DESIRED FUTURE: THE BRAIN INSTITUTE

As part of the Open Space for discussions at the very end of the workshop (see "Methods" on page 24), participants discussed desired futures and, notably, the creation of a European AI institute. We relay below the main points of discussion, which bring together many of the ideas from the previous sections.

Inspired by the European Organization for Nuclear Research (CERN) and the Massachusetts Institute of Technology Media Lab, the Brussels AI Network (BRAIN) Institute creates a hybrid environment that nurtures creativity and collaboration by connecting researchers from across disciplines, both the humanities and STEM. The Institute's primary aim is to research AI and ensure its robustness and the incorporation of European values, such as fairness in its developments. In addition, the institute facilitates the actual development and deployment of AI methods and models.

A NEW SCIENTIFIC CULTURE AROUND AI-ENABLED RESEARCH

Researchers at the institute work for short or extended periods, either at the central hub for collaboration and knowledge exchange, or at the satellite centres located across Europe that provide specialised expertise and resources. BRAIN combines a physical space with a robust digital infrastructure, recognising the importance of in-person interactions for nurturing creative ideas and fostering a vibrant research environment, creativity and team spirit.

The institute is financed through core funding, project-specific funding, and industry partnerships. This diversified approach ensures a sustainable financial base while enabling the institute to adapt to new research priorities and industry needs.

To encourage risk-taking and innovation, BRAIN promotes a 'fail is OK' culture, recognising that experimentation and setbacks are essential steps toward groundbreaking discoveries. This mentality encourages forward-thinking, which converts expertise and resources into tangible solutions for societal impact. Co-operation of researchers from diverse fields fosters cross-pollination of ideas and accelerates the development of groundbreaking AI applications.

BRAIN pursues a comprehensive theory of AI, aiming to go beyond its previous, practice-driven definition. By exploring the philosophy, ethics, and social implications of AI, the institute ensures that AI is developed responsibly and ethically. Instead of aiming to develop faster algorithms, BRAIN leverages humanities research on AI to explore how AI impacts research and its processes. It also focuses on explainable AI research, interacting with AI models, and using AI to solve complex global challenges.

ADAPTED INFRASTRUCTURE TO SUPPORT AI-ENABLED RESEARCH

BRAIN's digital infrastructure hosts the European Cloud Computing Centre. A collection of data is accessible to researchers across Europe through a centralised, secure cloud platform, offering a large, centralised repository of certified datasets for scientific use, ensuring the availability of high-quality data for AI development.

Sensitive data is handled with utmost care, ensuring compliance with EU data protection regulations, and employing robust anonymisation practices (thus playing a similar role to the existing European Molecular Biology Laboratory European Bioinformatics Institute which performs this role for omics and genetics data).

The cloud relies on a centralised EU graphics processing unit and includes open model repositories and well-curated, interoperable datasets, which promotes uniform standards for sharing and certification of data and models.

INFORMING POLICYMAKERS

The BRAIN approach simplifies data administration and accessibility while promoting a culture of data sharing and collaboration and accelerating AI research and development. BRAIN also invites and supports policymakers and engages with stakeholders from across sectors to bridge the gap between research and governance.

Implications for the European research and innovation ecosystem

Public-private partnerships

Participants discussed ways in which Europe could maintain its global relevance in AI advancement. One suggested way is to establish European Big Tech, with significant resources for data infrastructure and innovation. This could be achieved by engendering an entrepreneurial spirit in scientists and citizens from their early years.

The participants raised concerns about the potential misuse of AI by private actors. One way to prevent this is to champion the ethical use of AI through coalitions with the private sector and prevent nefarious uses of AI. For example, biological AI models could potentially be used to create dangerous chemical compounds.

In addition, at the interface of public and private science, the participants reflected that, in an era of AI, patents may need to be reconsidered, as data and model sharing could lead to the same results being obtained quickly in many places at once. Participants suggested that the notions of copyright and intellectual property may be rethought, taking into account who they benefit the most.

Geopolitics and security

The participants noted that if the current trends continue, AI will accumulate power in certain regions of the world and for certain sectors of society, thereby increasing inequality. Despite generating innovation, value, and growth, it can also be controlled by the military and nation-states and wielded in a race toward domination. It is interesting to note that the influence and reach of Big Tech can also surpass national boundaries and exercise a powerful voice equivalent to that of a nation. Even research institutes will undergo this process, with a few supersized ones leaving most others feeble and uncompetitive.

The participants highlighted that AI research at the present time is concentrated in a few digital companies, with monopolies in certain parts of the world. Some continents are completely left out of the equation. Participants reflected on how to act toward an equitable distribution of AI resources in the world by 2035. One idea brought forth at the workshop was championing widespread and open use of AI. Example initiatives include promoting open-source resources and tools (e.g., sandboxes), organising hacker fairs and hackathons, opening AI Maker Labs, giving libraries access to AI, and including AI theory and practice in school curriculums.

Participants also reflected that adverse effects should be monitored in providing AI access to everyone. A balance would need to be struck between openness and sharing on the one hand, and cybersecurity

vulnerabilities on the other. For example, an attack launched on the energy grid could paralyse an entire city. For this reason, it would be important to avoid having a single point of failure in our cyber systems.

AI infrastructure and environmental impacts

Participants discussed the implications of AI infrastructures for the environment. Existing centralised AI models put a great strain on our energy infrastructure. Their initial training and subsequent operation expend tremendous amounts of power. Participants reflected on the need to introduce measures, such as an AI tax, to offset the technology's energy use and its environmental impact.

In addition, participants also discussed the great need for raw materials used to manufacture computer chips, which constitute the building blocks of AI. The growth of AI might deplete those resources, which participants imagined could exacerbate inflation and lead to conflicts or wars. With perspectives of developing new and more sophisticated AI models by 2035, participants considered ways to ensure the sustainability of development that make use of finite and even scarce raw materials, ranging from finding alternative sources for materials to the enforcement of green AI implementation.

The implications for society

Participants brought up several times the potential implications of AI-enabled science on the society as a whole. These included liability, decision-making, work and AI-free alternatives. These reflections are summarised below.

Participants reflected upon a future in which, at some point, humanity might decide to confer legal identity on AI, just like it did on corporations. In that case, who should approve and legitimise decisions that affect AI? Or would AI govern itself, drafting and implementing legislation, or even voting? Could AI become self-serving? For instance, it could direct resources to itself and ultimately threaten the existence of humanity.

Whether AI is vested with legal identity or not, the question remains: who is liable for its decisions, including its wrongdoings and misconduct? AI models often exhibit reasoning on par with that of humans.[6] However, decisions made by AI are data-driven, and not value-driven (except in how these values are encoded in the data). This is an important difference with humans making decisions.

Adding complexity to this question, participants also reflected upon understandability. AI is able to quickly generate solutions that work, but AI methods are unable to explain why they give this solution. The current understanding of AI models, especially in the general sense, is low. In the future, there is a need to explore the discrepancy between the understanding of AI outputs and the effects they have on our lives.

6 *AI Has Evolved to Reason Like Humans, Scientists Say*, Darren Orf, Popular Mechanics (2023), about about arXiv:2303.12712v5. https://www.popularmechanics.com/technology/robots/a43906996/artificial-intelligence-shows-signs-of-human-reasoning/

Not only scientists, but society as a whole may be impacted by the developments of AI. Participants discussed a future in which job displacement and unemployment will affect many professions, like journalists, writers, and even scientists. In an imagined scenario, this could reduce working hours or completely eliminate work in favour of Universal Basic Income. However, history tells us it is unlikely. The washing machine is a good example. Essentially an early example of automation, it was expected to free humans of labour. Instead, it gave them space to do other chores. Similarly, AI might give rise to new professions, such as AI maintenance workers and AI facilitators, who will interact with AI to keep humans in the loop.

Finally, participants discussed the importance of ensuring freedom with respect to AI use by Europeans. Not everyone will be able or willing to use AI. Future societies may need to accommodate AI sceptics by offering AI-free activities and spaces, or alternative options (such as word processors that do not use AI in the backend). Furthermore, there should be an awareness that an analogue space can still be impacted by AI, for instance when social media algorithms make a particular location well-known and direct the tourist stream there. Thus, AI-free should be considered in a specific (free from direct AI use) and more general (indirectly free from AI) sense.

Governance insights

Alongside the workshop and foresight exercises, the participants provided governance insights and options for the EU to consider. The main ones are listed below.

For data and models:

- defining boundaries of AI-made works (what is considered AI-made versus AI-contributed) and enforcing a watermark system to label AI-made work
- developing a way to trace and label data with respect to its origin
- establishing a certification system for AI training datasets
- ensuring that AI training datasets conform to the FAIR principles
- championing the equivalent of fair trade for data training, regulating how the data is trained, by whom, and under what conditions
- setting up a Cloud Computing Centre in Europe, furnishing it with AI models and large quantities of datasets

For science and education:

- Preventing inequality in education, and therefore providing access to AI, by promoting and funding skill training initiatives accessible to all.
- Promoting AI literacy and capacity building with scientists and citizen representatives at the helm.
- Offering long-term (decades-long) funding for academic research.
- Encouraging transdisciplinary science initiatives.

For the EU research and innovation ecosystem:

- prioritising sustainability in AI and pushing for green AI initiatives
- encouraging collaboration between the public and private spheres by creating opportunities and incentives
- championing the ethical use of AI through coalitions with the private sector
- acting toward an equitable distribution of AI resources in the world

For society:

- providing AI free/alternative spaces
- preventing the use of AI to enforce unrealistic and inhumane standards of productivity and efficiency on humans, e.g., through surveillance methods such as eye-tracking
- championing the defence of privacy, providing incentives, involving social scientists and decision-makers in the process, and leveraging industry partnerships

Conclusion: how can AI boost scientific productivity?

AI will impact science in a variety of ways. It will shift emphasis from scientific publications to an emphasis on the production, management and use of data; assist in review and assessment processes; allow for new career options; and redefine a typical 'workday' schedule. It can help us acknowledge and investigate bias, make us rethink knowledge ownership, and importantly, assist scientists with the mundane tasks such as paperwork and empower them to become natural philosophers once again, hinting at a new age of enlightenment.

On the other hand, it is important to keep in mind the possible disadvantages. Although the process of training AI models on data is based on solid mathematics and statistics, its outputs cannot be given a logical explanation and therefore understood, giving the users the impression of an oracle. Moreover, often the big tech AI companies do not make their data conform to the FAIR principles and their methods open to public scrutiny. AI may challenge our economic systems and has the power to reconfigure geopolitical heavyweights. In terms of doing science, relying on it as a production tool might remove serendipity from the scientific endeavour.

Tracing the future trajectory of AI is rife with uncertainty, we are at a decisive time when we can ensure the implementation of AI that champions the values of the EU, while bolstering global competitiveness and independence.

Annexes

Story boxes

As part of the foresight workshop process, participants were asked to imagine the future of AI in science through stories and elaborate reflections. The story boxes below gather these imaginary scenarios and serve to illustrate some important points made in the report.

STORY BOX: HOW MUCH CAN WE RELY ON AI?

CASE 1: REFLECTIONS ON DIGITAL TWINS

In the context of scientific discovery, the term "uncertainty" is positive. In fact, uncertainty enables discovery. In domains of high complexity, AI tools can provide an integrated understanding of the subject — notably through the development of so-called 'digital twins' (digital replicas of physical, virtual or hybrid objects). This approach is promising in medicine, social sciences, epidemiology, urban systems, and climate change research — as in the case of Earth Digital Twin, an EU programme that employs digital twins to explore the effects of climate change .

However, a digital twin is tied to something that already exists — an old view of the world. It gives us a sense of control while removing an element of chance. These types of models require huge amounts of open access resources to be built and to stay relevant. Relying on AI for policymaking could take evidence-based decision-making to the extreme.

CASE 2: IMAGINED FUTURE - S.A.M. SOLVES CLIMATE CHANGE

In 2035, Earth is a borderless place where interdisciplinarity reigns supreme. The experts are polymaths, educated through grants from transdisciplinary projects. They possess a wide range of knowledge in various fields, make connections between different areas, and use mutually understandable terminology.

This society functions similarly to the contemporary game industry, where several professionals from different domains work together toward a common goal.

In this environment, human committees of experts supervise AIs in solving complex problems and navigating between dystopian and utopian outcomes. One of those AIs is S.A.M., short for Synthetic Articulator Moderator, who succeeds in solving climate change and goes on to enable scientists to reach unexpected conclusions on different topics, becoming a catalyst for change and a facilitator across different fields.

STORY BOX: TRANSHUMANISM

WHO ARE WE?

Transhumanism touches on the mind-body problem, which asks the question of whether having a physical body is crucial for the existence of consciousness. Although it may not need to be the entire body. Neuroscientists emphasise the importance of the brain, suggesting it must remain intact even as all the rest gets replaced.

Theoretically speaking, sufficiently advanced technology could allow people to move their consciousness to the cloud. They could choose that option after having lived in bodies until retirement. However, would the person in the body be the same person as the person in the cloud?

The Ship of Theseus is a thought experiment that explores the conundrum of identity and self. If we replace all the elements in an object with new ones, but take care to preserve its exact structure and composition, can we say it is the same object as before or not? The same idea applies to the human body. For this reason, the notion of placing consciousness is problematic. To begin with, is it even possible? Are we still ourselves if we abandon our physical form? Or does our body inform the brain and shape aspects of our personality and perception?

SUSTAINING THE BODY

Another possibility is to maintain our physical form for a longer time. One might argue this is already happening. Modern advances in medicine have extended the human lifespan so much that it becomes a challenge for policymakers, e.g., testing the limits of our pension schemes.

STORY BOX: A SCIENTIST OF TOMORROW

PURISTS VS ADOPTERS

In 2035, most EU citizens will be educated in AI. They will enjoy AI-free leisure and activities in designated areas. Regulation will prevent AI monopolies, allowing diverse ideas to flourish. At the same time, the ubiquity of AI will not dissuade some people (called Purists) from shunning it. They will rely on their natural intelligence and look down on AI Adopters.

Perhaps with good reason. The Adopters' overreliance on AI will hamper their cognitive abilities and make them vulnerable to manipulations such as false information injection. Compromised scientists will try to save faces by performing "AI detoxes".

AI misinformation will continue to affect entire communities and domains—from voting to climate change awareness to general science. Citizens Assembly and high-ranking authorities will intervene to vet large datasets and ensure information authenticity.

SCIENTISTS AND ROBOTS

In 2035, most people, including scientists, are happily unemployed — or rather, job-free. Universal Basic Income satisfies their basic needs. All the science is done by self-sufficient robots doubling as laboratories. The robots not only do research but also publish their own papers in AI scientific journals.

Occasionally, a human supervisor finds fault in a robot's research — for example, plagiarism. The robot is tried before a human committee and decommissioned if found guilty. Although it loses its physical form and abilities, it remains active as software in the cloud.

List of participants and other attendees

Experts

- Jan Aerts, KU Leuven, Belgium
- Pelin Angin, Middle East Technical University, Turkey
- Stefaan Callens, KU Leuven
- Adel Daoud, Linköping University, Sweden and Chalmers University of Technology, USA
- Mathieu Denis, ISC Centre for Science Futures, France
- Nausikaä El-Mecky, Pompeu Fabra University, Spain
- Elle Farrell-Kingsley, Mother Tongue, UK
- Jaroslav Hlinka, Czech Academy of Sciences, Czech Republic
- Marijn Hoijtink, University of Antwerp, Belgium
- Dirk Hovy, Bocconi University, Italy
- Gregor Kasieczka, Hamburg University, Germany
- Gražina Korvel, Vilnius University, Lithuania
- Lydia Koustopolous, LKCYBER, UK
- Jaakko Kuosmanen, Finnish Academy of Science and Letters, Finland
- Lenka Lhotská, Czech Technical University in Prague, Czech Republic
- Ewa Luger, University of Edinburgh, UK
- Benoît Macq, Université catholique de Louvain, Belgium
- Jan Madsen, Technical University of Denmark, Denmark
- Pegah Maham, Stiftung Neue Verantwortung, Germany
- Sébastien Massart, 3DS Dassault Systèmes, France
- Marc Mézard, Bocconi University in Milan, Italy
- Roman Neruda, Czech Academy of Sciences, Czech Republic
- Max Reddel, The International Center for Future Generations, Belgium
- Philip Shapira, University of Manchester, UK
- Andrzej M.J. Skulimowski, AGH University of Science and Technology in Kraków, Poland
- Joos Vandewalle, KU Leuven, Belgium
- Lidia Zuin, Disruptive Futures Institute / Techistential, Sweden

School of International Futures

- Peter Glenday, Programme and Research Director
- Darja Vrščaj, Foresight Researcher and Project Manager

SAPEA Board and staff

- Stefan Constantinescu, Chair of the SAPEA Board
- Marie Franquin, Scientific Policy Officer
- Stephany Mazon, Scientific Policy Officer
- Rúben Castro, Scientific Policy Officer

Science writer

- Hubert Brychczyński

Science Policy, Advice and Ethics Unit at DG RTD, European Commission

- Ingrid Zegers, Team Lead
- Jean-Francois Dechamp, Policy Officer
- Daniela Melandri, Policy Officer
- Gintare Juskaite, Policy Officer

Observers from the European Commission

- David Arranz, Policy Officer, RTD.E4
- Susana Nascimiento, Coordinator and Adviser for Ethics Review and Monitoring, ERCEA B0
- Daniela Petkova, Policy Officer, RTD.E4
- Lisa Wetzig, Blue Book Trainee at the European Commission, RTD.E4

Methods

Workshop approach

In response to the request for input from Margrethe Vestager, Executive Vice-President of the European Commission, to the Group of Chief Scientific Advisors on how the European Commission can accelerate a responsible uptake of AI in science a workshop was organised by SAPEA on 26–27 October 2023 to address key area 1 of the scoping paper, on vision and foresight.[1]

The workshop addressed the question:

> What impetus could Artificial Intelligence give to scientific productivity and what benefits, incentives and challenges would AI-enabled research bring to the European innovation ecosystem and society as a whole?

The workshop was designed and facilitated by the School of International Futures based on the questions set out in the Scoping Paper, with support from SAPEA and the SAM.

It used a strategic foresight approach[2] — an organised, systematic process for engaging with uncertainty and looking beyond the expected — to generate ideas, debate key issues and imagine possibilities for AI-enabled research in 2035. It then looked to identify key benefits, incentives and challenges. The aim of the workshop was not to predict the future, but to understand what is changing and what this could mean in the future. The design of the workshop was participatory and exploratory, building on the experience and knowledge of the participants.

Participant selection

In the context of the request for advice on this topic, SAPEA issued a call for nominations on 18 July 2023, describing the scope, timeline and expertise required. The call for nominations was sent via the academy networks to their member academies, who were invited to nominate experts following the procedure described in the SAPEA Quality Assurance Guidelines.[3]

The list of areas of expertise that should be covered in the foresight workshop was established in coordination with the Group of Chief Scientific Advisors, the SAM secretariat and SOIF. This list included AI governance, policy and technologists, science policy and grand challenges, arts and humanities, health sciences, STEM and sustainability, cybersecurity and defence, as well as futures thinkers with expertise in science and technology.

1 Key areas 2, 3 and 4 of the Scoping Paper were addressed separately by SAPEA, through a process of evidence review supported by a working group.

2 European Commission (no date). *Strategic foresight*. https://commission.europa.eu/strategy-and-policy/strategic-planning/strategic-foresight_en

3 SAPEA. (2023). *Quality assurance guidelines and Procedures on science advice for policy and society*. Berlin: SAPEA. https://doi.org/10.5281/zenodo.8329539

Experts participating in the foresight workshop were selected from the resulting list of nominees. Additional experts were also identified through desk research by the academy networks, suggested by the SAM secretariat, European Research Council Executive Agency and the SOIF.

The experts were selected by SAPEA on the basis of scientific excellence and disciplinary requirements as a priority, taking into account commitment and time availability, and the criteria set out in our Strategy of Diversity and Inclusiveness:

- interdisciplinarity and multidisciplinarity
- involvement in the wider scientific community
- inclusion of early- and mid-career researchers
- balanced gender representation
- wide geographical coverage, including from Widening countries

The list of selected experts was approved by the SAPEA Board on 3 October 2023, prior to sending invitations.

The final group of experts who attended the workshop included 54% male and 36% female experts. 18% of attending experts were early-career and 43% mid-career. 21% of experts came from Widening countries. In total, the following countries were represented: Belgium, Czechia, Denmark, Finland, France, Germany, Italy, Lithuania, Poland, Spain, Sweden, Turkey, and the United Kingdom.

In line with the SAPEA principle of transparency, workshop expert participants were asked to declare any conflict of interest and any interest that might be perceived by SAPEA as a conflict of interest in relation to this scientific topic. The information declared by experts was made available to the workshop participants on 26–27 October 2023 in printed form at the premises of the European Commission.

Workshop format

The Chatham House rule applied to the entire workshop. In advance, participants had received a copy of the scoping paper, a short introduction to foresight, and a draft copy of a SAPEA literature review on *Successful and timely uptake of artificial intelligence in science in the EU*.

After a general introduction to the science advice process, participants were introduced to the format and agenda for the two days. The key steps in the workshop were:

1. **Identifying key themes.** Participants used a found postcards[4] exercise to identify key topics that are not getting enough attention in relation to AI-enabled Science, Research and innovation in 2035. These ideas were shared and refined through a series of rotations to identify nine topics.

4 This method uses found, as opposed to purposefully selected, postcards as a particular type of a found object, as a device to enable dialogue in groups that allows participants to build new stories about the future. Curry A., and Ward V., (2014) *Postcards as doorways*. Journal of Futures Studies, March 2014, 18(3): 101-114. https://jfsdigital.org/wp-content/uploads/2014/04/183-E02.pdf

These were clustered into five thematic areas. Each thematic area was then explored on by one group.

2. **Developing scenario vignettes.** Participants were supported to create scenario vignettes that described one possible future for Science Research and innovation in 2035. To do this, they used a series of futures methods that help develop the logic and narrative of future worlds. The key steps in the process were:
   - using futures wheels,[5] a brainstorming tool that helps identify second- and third-order impacts of changes, to identify important and/or transformative changes that might emerge in each of the thematic areas
   - cross-impact analysis,[6] to identify how key impacts might combine to impact science, research and innovation in 2035
   - using a combination of structural questions, ethnographic[7] and narrative futures[8] techniques to build rich pictures of the world in 2035. These were produced in the form of a cover page from a popular science magazine

3. **Imagining change.** Having developed their scenario vignettes, the three horizons[9] technique was used to help participants to identify key benefits, incentives and challenges associated with each of these futures. This system change model allows for an intuitive understanding of how futures might emerge, and provides an understanding of opportunities, obstacles, and gaps which exist in the emerging futures landscape. Using the technique, participants considered the scenario vignettes (horizon 3), the present (horizon 1) and the transition (horizon 2). They were provided with questions from the scoping paper to stimulate the group discussions. Finally, they were then asked 'what do we need to do in the transition to ensure the responsible uptake of AI in science, research and innovation?'.

4. **Open space: group discussions.** The next step in the workshop was to provide an Open Space[10] for participants to host conversations on topics that they chose. This was an opportunity to focus in

5 The Futures Wheel technique was invented by Jerome Glenn. Glenn J., (2021) Futures Research Methodology - v3.0, published by The Millennium Project. The technique was adapted for this workshop to allow for explicit exploration of social, technological, economic, environmental, political and legal impacts, and values. https://jeasprc.org/wp-content/uploads/2020/08/06-Futures-Wheel.pdf

6 We used a light-touch cross impact approach drawing on the Manoa Method that asked participants to imagine how three impacts might combine to shape the future. Curry, A, Schultz, W (2009). *Roads less travelled: different methods, different futures.* Journal of Futures Studies, May 2009, 13(4): 35–60. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fc73132a09cb7be8051427c7dd1db74aa8fb6607

7 In particular, we adapted the Verge (Ethnographic Futures Framework) developed by Michele Bowman and Richard Lum: Lum R., (2014). *Working with Verge*. APF Compass, April 2014. https://ddtconference.files.wordpress.com/2017/07/lum-verge-apfcompass-april14.pdf

8 The approach drew on character-led and narrative futures techniques developed by Emily Spiers. Lively, G, Slocombe W., Spiers, E (2021). *Futures literacy through narrative*. Futures, 125, 102663 https://www.sciencedirect.com/science/article/abs/pii/S0016328720301531?via%3Dihub

9 The Three Horizons approach was developed by Anthony Hodgson, Bill Sharpe, and Andrew Curry. Curry A., and Hodgson A., (2008). *Seeing in multiple horizons: Connecting futures to strategy*. Journal of Futures Studies, August 2008, 13(1):1-20. https://www.internationalfuturesforum.com/3hblog/wp-content/uploads/2008/11/3h-curry-hodgson-jfs1.pdf

10 Invented by Michael Herman. Herman M. (2016) "Inviting Leadership in Open Space. A Guide for Training and Practice." https://michaelherman.com/publications/inviting_leadership_guide.pdf

on issues that had emerged from across the scenario vignettes with a view to achieving a positive future for AI-enabled science research and innovation.

5. **Key takeaways.** A final exercise asked participants to share their key takeaways and identify key priorities for the European Innovation Ecosystem.

Processing workshop outputs and reporting

All workshop outputs were captured on post-its and flipcharts. One note-taker was present at each table of 5–6 participants during the workshop, and all group and plenary discussions were recorded and transcribed where necessary.

The transcriptions and notes were collated in one document and a sense-making and coding approach used to identify key themes for inclusion in the report. For the report writing, the notes, post-its and flipchart content were summarised into the three horizons for each group of experts based on the collated notes document. Key themes were identified by SAPEA and SOIF based on the main question of the workshop and the content of discussions. The contents of the three horizons for all tables were colour-coded based on the main themes. The present report was then written to summarise the content under each theme.

Programme

26 October 2023	
09:30	Welcome and brief topic introduction by SAPEA and European Commission
09:45	Introduction, ways of working and icebreaker
10:05	Introduction to foresight
10:20	Surfacing implicit knowledge with the found future postcards method
11:45	Identifying patterns of change
12:15	Lunch
13:15	Exploring impacts of change
14:45	Plenary discussion
15:30	Break
15:45	Building worlds of change
17:20	Reflections and close
27 October 2023	
09:00	Welcome
09:15	Identifying opportunities, gaps and obstacles with the Three Horizons method
10:45	Break
11:00	Open Space
11:50	Plenary discussion
12:15	Key takeaways and priorities
12:25	Close

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).