

Frontier to forefront: the ascent of AI in modern research

The acceleration of progress

From its beginnings in the 1950s, artificial intelligence (AI) has always been a groundbreaking field at the frontier of research. In recent years, a slow march to realise true AI has been replaced by rapid advances due in large part to the emergence of generative artificial intelligence (GenAI).

A 2023 survey by [Nature](#) asked more than 1,600 researchers around the world how useful they thought AI tools would become for their fields in the next decade. More than half of those surveyed expected AI tools to be 'very important' or 'essential'.

On 21 March 2024, the United Nations General Assembly adopted a United States-led [resolution](#) on the promotion of 'safe, secure and trustworthy' AI systems that will also benefit sustainable development for all. Backed by more than 120 other Member States, it represents the first time the Assembly has adopted a resolution on regulating AI and recognises 'the rapid acceleration of the design, development, deployment and use of artificial intelligence systems and rapid technological change'.

The catalyst

GenAI refers to a type of AI capable of generating original content, such as text, images, audio or even entire pieces of multimedia not directly copied from existing data. First introduced by OpenAI in June 2018, the Generative Pre-trained Transformer (GPT) architecture was made public with GPT-1, a large-scale unsupervised language model.

The most recent version is GPT-4, released on 14 March 2023. It can generate, edit and iterate with users on creative and technical writing tasks. Capable of processing over 25,000 words of text, OpenAI cites use cases including long form content creation, extended conversations, and document search and analysis.

Crucially, these recent advances have pushed AI beyond the laboratory into the wider research sphere. For research managers, it has the potential to transform how tasks are carried out, such as processing funding applications, monitoring project progress and managing financial reporting.

AI tools are already used for a manner of different tasks within the research sector. Researchers at the Institute for Methods Innovation at the University of Warwick recently published an [article](#) on the way researchers may be able to utilise GenAI tools like ChatGPT to increase productivity.

Suggestions include the creation of more accessible science communication content for different audiences and contexts, and repurposing funding applications to create impact plans or adjusting an application to fit a particular call. The tools could also help with ideas for survey design to evidence research impact or support the analysis of survey data. The researchers also highlight the tools' limitations and note that they are 'great support tools' intended as 'a starting point for content generation, editing and improvements'.

Action and reaction

In December 2023, a European Commission [policy brief](#) on the role and potential of AI in science and innovation advocated a tailored ERA (European Research Area) policy to speed up and facilitate the adoption of AI in science. This would harness the potential of the technology to accelerate breakthroughs, foster innovation and drive economic growth. Included in the brief are also infrastructural aims such as strengthening of enabling data and computer ecosystems.

At the same time, the Commission acknowledges AI-related ethical challenges and highlights the need to monitor the impact of AI on careers in research as well as on public trust in science in general.

Outside the EU, the UK is seeking to establish a leadership role within a broader internationalist approach. It hosted the world's first global [AI Safety Summit](#) in November 2023, bringing together leaders from 28 countries and the EU as well as leading players in industry, academia and civil society. The summit concluded with a Declaration for nations to work together to identify, measure and regulate the risks posed by AI.

The second edition of the summit was co-hosted by the UK government in Seoul, South Korea, over 21-22 May 2024, which brought together international governments and leaders from global industry, academia and civil society. One of the key outcomes was the [Seoul Declaration for Safe, Innovative and Inclusive AI](#), which underscores a commitment to engage with various global initiatives and foster international cooperation to help develop AI governance frameworks. France is set to host the third AI Safety Summit later in 2024.

During the initial summit, the UK announced the formation of the [AI Safety Institute](#), the first state-backed organisation dedicated to advanced AI safety for the public interest. The new body will make its findings available worldwide to facilitate an effective global response to the opportunities and risks posed by advanced AI. The Institute's [Systemic AI Safety Fast Grants](#) will enable researchers to explore how to protect society from AI risks such as deepfakes and cyberattacks as well as helping to harness its benefits such as increased productivity. The most promising proposals will be developed into longer-term projects and could receive further funding, with the total value amounting to £8.5 million.

Other funding opportunities designed to enhance the safety of AI are offered through [Responsible AI UK](#), a £31 million international ecosystem for responsible AI research and innovation. Funds include a skills programme to develop open educational resources for the training of AI; an impact accelerator to help ensure AI technologies are designed, deployed and used responsibly within societies; and international partnerships to foster collaborative research on responsible and trustworthy AI. Elsewhere, [Bridging Responsible AI Divides UK](#) represents a £15.9 million investment designed to embed arts and humanities research within the responsible AI ecosystem while bridging divides between academic, industry, policy and regulatory work on responsible AI.

These risk management measures are allied with a significant drive to promote and capture the benefits of AI. In February 2024, the UK invested £80 million in [nine research hubs](#) that will deliver next-generation innovations and technologies in areas such as supporting better health treatments and delivering faster development of electronic devices and microchips. Work is also underway to establish a new [AI Research Resource \(AIRR\)](#), the UK's fastest supercomputer capable of making 200 quadrillion calculations a second. Dubbed Isambard-AI and based at the University of Bristol, the resource is scheduled to open in 2024 and aims to put the UK at the forefront of AI research and innovation.

Elsewhere, UK Research and Innovation's (UKRI's) [five-year review of The Alan Turing Institute](#) – the UK's national institute for data science and AI – concluded that core public funding of a minimum of £15 million per year should be allocated to the institute provided that certain conditions are met, including a revised governance structure and a clear financial plan covering the next five years. Additionally, an AI map of the UK known as [WAIFinder](#) has been launched to enable researchers and innovators to explore the breadth and range of the UK's research and innovation ecosystem. This free-to-access system is constantly updated and allows researchers and innovators to dynamically browse companies, funders, incubators and academic institutions that are involved in creating AI products, services, processes and research.

Such initiatives follow the spirit of the UK government's March 2023 [AI Regulation white paper](#), which set out a 'pro-innovation' approach to regulation. This approach consciously aims to avoid regulation at the early stages of AI's development before the risks are properly understood. A February 2024 [response](#) to a consultation on the initial white paper stated that the UK government will 'not rush' to legislate and 'will legislate when we are confident that is the right thing to do'.

The global reach of the technology has led to a consensus that international collaboration is essential. On 1 April 2024, the US and UK signed a [memorandum of understanding](#) on developing AI safety, with a view to further increase alignment over time. Measures include committing to perform at least one joint testing exercise on a publicly accessible model; collaborating on AI safety technical research; fostering personnel exchanges; and sharing information about activities conducted. The memorandum also contains a commitment to working with other governments on international standards for AI safety testing, as well as on other relevant standards.

This commitment was swiftly realised on the US side following the sixth EU-US Trade and Technology Council (TTC) meeting in April 2024, which saw the creation of a new research alliance called '[AI for Public Good](#)'. This partnership seeks to ensure EU and US scientists work collaboratively on innovations that deliver societal wellbeing. Recent developments in areas including energy optimisation, emergency response, urban reconstruction, and extreme weather and climate forecasting will be built upon as the partnership seeks to forge further collaborations and address global challenges.

At the same meeting, 13 new terms and 24 amended terms were included in the [EU-US Terminology and Taxonomy for Artificial Intelligence](#) document, which aims to provide a standardised approach to AI risk management and trustworthy AI in the EU and US. New terms include 'deep fake' and 'confabulation' (aka hallucination) while amended definitions were made to terms including 'large language model', 'trustworthy AI', 'natural language processing' and 'big data'.



The research response

While governments attempt to address the growing influence of AI on national policy, the international research community faces the impact of rapidly evolving AI on traditional working practices while ensuring the continued integrity and validity of research.

The **European University Association (EUA)**, which represents more than 870 universities and national rectors' conferences in 49 European countries, set out a **position paper** on the responsible use of AI tools in higher education teaching and learning, in which it described any attempts to ban the technologies as 'futile'.

It instead advises institutions to adapt their approaches to the reality of the technology and invites universities to consider key factors related to AI tools' immediate consequences on learning and teaching, the ongoing innovation of learning and teaching, and universities' broader role in relation to the impact of such technologies on society.

The UK Russell Group of 24 leading universities took a similar approach. A July 2023 **statement** acknowledged the risks and benefits of AI technologies and set out five principles to ensure staff and students are 'AI literate' and able to capitalise on the opportunities that AI technologies can provide for teaching and learning:

- Universities will support students and staff to become AI-literate.
- Staff should be equipped to support students to use GenAI tools effectively and appropriately in their learning experience.
- Universities will adapt teaching and assessment to incorporate the ethical use of GenAI and support equal access.
- Universities will ensure academic rigour and integrity is upheld.
- Universities will work collaboratively to share best practice as the technology and its application in education evolves.

This need to foster 'AI literacy' was subsequently identified as one of the key takeaways of a November 2023 **conference** organised by the Coimbra group of 41 European universities exploring the use of GenAI in universities, which identified potential negative impacts on information integrity but nevertheless emphasised the 'importance of maintaining calm'.



Dr Newman-Griffis from the University of Sheffield in their **AI thinking and the enterprise of science** preprint speculates that approaching AI as a methodology rather than a technology can help manage the tension between promise and peril. They propose an 'AI thinking' perspective – the process of how we design and use AI systems to solve problems in specific contexts – can help scientists achieve richer analysis, more open science and constructive use of GenAI while helping manage risk of harm: 'Just as statistical thinking transformed how scientists defined and worked with evidence in nineteenth and twentieth-century science, so AI thinking can help transform how scientists work with data in the twenty-first century. Science has always required putting tools and evidence in context, and it is exactly this process-oriented perspective that must guide our understanding and use of AI in science.'

The European Commission together with the ERA countries and stakeholders published a set of **guidelines** on 20 March 2024 to help the European research community – including researchers, research organisations and research funders – use GenAI responsibly. The guidelines aim to ensure that a coherent approach to the technologies applies across Europe and are designed to balance the technologies' benefits with their potential risks to maintain research integrity.

Key takeaways from the guidelines include:

- Researchers should refrain from using GenAI tools in sensitive activities such as peer reviews or evaluations and use GenAI respecting privacy, confidentiality and intellectual property rights.
- Research organisations should facilitate the responsible use of GenAI and actively monitor how these tools are developed and used within their organisations.
- Funding organisations should support applicants in using GenAI transparently.

The guidelines are not binding and instead take into account pre-existing frameworks for the use of AI both generally and in the research sector specifically, such as the **European Code of Conduct for Research Integrity** and the **guidelines** on trustworthy AI. Users of the new guidelines are encouraged to adapt them to their specific contexts and situations, keeping proportionality in mind.

Given the rapidly evolving nature of the technology, the guidelines will be updated with regular feedback from the scientific community and stakeholders. Interested stakeholders are therefore invited to submit their feedback on the guidelines via the **dedicated portal**.



AI and funders

The pragmatic approach to the technology taken by the academic sector is mirrored by research funders. In September 2023, a [joint statement](#) was issued by the UK Research Funders Policy Group, a collective of senior representatives from some of the UK's major science and health research funders, including UKRI, the Association of Medical Research Charities (AMRC), Cancer Research UK (CRUK), the National Institute for Health and Care Research (NIHR) and the Wellcome Trust.

While recognising the opportunities and benefits of GenAI tools, such as assisting neurodivergent researchers and reducing potential language barriers, the statement highlights multiple risks in both the application and assessment process of funding applications, including relating to 'rigour, transparency, originality, reliability, data protection, confidentiality, intellectual property, copyright, and bias'.

Regarding the development of funding proposals, the guidance states researchers 'must ensure generative AI tools are used responsibly and in accordance with relevant legal and ethical standards where these exist or as they develop'. The statement also forbids peer reviewers from using GenAI tools in their peer review critiques or applicant responses to critiques, citing confidentiality concerns.

In the same month, Deutsche Forschungsgemeinschaft (DFG – German Research Foundation) issued a similar statement discussing the influence of generative models of text and image creation. The [statement](#), which was also [endorsed by Volkswagen Foundation](#), offers initial guidelines for dealing with GenAI tools in the context of scientific and scholarly research, as well as within the framework of the review, evaluation and decision-making process of research funders.

As per its UK counterparts, DFG does not rule out the use of generative tools but highlights the need to uphold certain standards to ensure good research practice. This includes aspects such as transparency and verifiability of research processes and findings, compliance with professional ethics and basic principles of research integrity, as well as disclosure of all cases where generative models are used.

The statement highlights the responsibility of the human author of any research publication to ensure that the use of AI tools does not infringe other people's intellectual property or result in scientific misconduct. DFG also makes clear that the use of generative models in the preparation of research proposals submitted for funding has no impact on the funding decision. However, the same tools cannot be used for the preparation of reviews as this may breach the confidentiality of the process.

Safeguarding and looking ahead

A recent [LSE blog post](#) on the topic posits that recent developments are part of a long-term trend, stating: 'Claims to moral purity based on avoiding GenAI tools are dubious at best. Anybody who has ever accepted a suggested change from an automated spelling or grammar check in Microsoft Word already has some experience with using an AI tool to improve their work.'

However, ethical concerns outlined in statements made by the Research Funders Policy Group and DFG related to the rapid expansion of GenAI and the need to manage risk are being raised across the research community.

On 13 March 2024, the European Parliament approved the [Artificial Intelligence Act](#) (AI Act). It aims to 'protect fundamental rights, democracy, the rule of law and environmental sustainability from high-risk AI, while boosting innovation and establishing Europe as a leader in the field'.

The AI Act introduces rules banning certain AI applications that threaten citizens' rights. Emotion recognition in the workplace and schools, social scoring, predictive policing and AI that manipulates human behaviour or exploits people's vulnerabilities will also be forbidden.

In concert with the AI Act, the European Commission's new European AI Office will play a key role in implementing its provisions and fostering the development and use of trustworthy AI.

These developments run in parallel with concerted efforts to ensure Europe does not fall behind in the race to be at the forefront of AI development. Crucial to this will be the preparation of the successor to Horizon Europe, Europe's Framework Programme for Research and Innovation. Currently referred to as FP10, the Commission is scheduled to publish a set of proposals in July 2025, offering a glimpse into its long-term plans for meeting this dual challenge.

However, with AI tools developing at an exponential rate, specific plans can quickly become outdated and less important than setting out how long-term challenges will be faced. FP10 will need the flexibility and adaptability necessary to meet whatever challenges lie ahead and, crucially for many in the sector, the large-scale investment needed to ensure Europe can compete in this most competitive and crucial field.

The 2024 European R&I Days highlighted the need to foster strong public-private partnerships to both enable sharing of the extensive and expensive resources required to conduct this research, and to ensure that the vast amount of data and discovery research conducted in the public realm can be successfully translated to deliver the specific needs of industry. President of the European Partnership in AI, Data and Robotics (ADR), Emanuela Girardi, highlighted that the Partnership had been most successful when it moved away from a generalised approach to AI and instead focused on the specific needs expressed by industry, such as the use of AI for robotics.

Addressing the major technological and logistical challenges facing AI over the next 10 years, Head of AI for Science at Google DeepMind, Dr Pushmeet Kohli, outlined both the need to rescale AI models as their capabilities progress and, crucially, to ensure scientists have access to them and are genuinely capable of understanding them. This point was backed by Girardi, who stated that the shift to AI requires genuine investment in the education of scientists so they can utilise AI technologies in a responsible way.





At the AI and research: a promising relationship? webinar event in May 2024 (organised by Research on Research (RoR) Registry and Hub, University of Southampton), Amanda Blatch-Jones, part of the RoR team, expressed the need for greater collaboration between the research sector and the private tech companies dominating the AI field, stating: 'Lack of understanding of the value, impact and sustainability of AI adoption remains a significant gap in the evidence. There is need for evaluation studies to understand whether AI technologies and AI-enabled innovations can achieve the intended goals, and help to build trust in the use, applicability and adoption for the application of AI.'

Disruptive and transformative, but loaded with risk and ethical concerns, harnessing AI's potential represents a huge challenge for the international research community.

As part of its ongoing coverage of research funding and policy developments, ResearchConnect covers breakthroughs in AI in and for the research sector. We offer a user-friendly database containing a global source of research opportunities, covering a broad range of funders and disciplines. Our news content features the most recent research calls and funder updates, providing access to the latest funding opportunities and developments to researchers worldwide.

Call us now on 0333 011 1200 or email marketing@idoxgroup.com for more information about Idox's funding services.

Idox Software Ltd
Unit 5, Woking 8
Forsyth Road, Woking
Surrey GU21 5SB

T: +44 (0) 333 011 1200
E: marketing@idoxgroup.com
www.idoxgroup.com

© Idox plc. 2024 Ideas, solutions, suggestions, hints and procedures from this document are the intellectual property of Idox plc and thus protected by copyright. They may not be reproduced, transmitted to third parties or used in any form for commercial purposes without the express permission of Idox.

