

An inflection point for scientific research

Summary of the 2024 US-UK Scientific Forum *Science in the age of AI*
held on 11 – 12 June 2024

Preface

The rise of artificial intelligence (AI) represents one of the most profound technological revolutions of our time, impacting on how we think, learn, and innovate across nearly every section of society. In science, we have witnessed the transformative impact of AI, enabling researchers to grapple with hitherto intractable problems, and driving progress in fields ranging from genomics to climate modelling to materials science. As the science behind AI grows ever more sophisticated, it has great potential to unlock even more possibilities. Yet these opportunities also present critical challenges. Addressing questions around the transparency, reproducibility, ethical use of and equitable access to AI technologies remains fundamental to its successful integration into scientific practice.

The 2024 US-UK Scientific Forum, *Science in the age of AI*, was held at the Royal Society, in London, on 11 – 12 June 2024. The forum brought together leading researchers and other key stakeholders from across the UK and the US to explore how AI is transforming the landscape of scientific research. The Forum provided a unique platform for interdisciplinary discussion, examining AI's potential to advance open science, drive groundbreaking discoveries, and address global challenges. Equally, it encouraged crucial discussions about the responsibilities scientists bear in harnessing this transformative technology while upholding the principles of rigor, integrity, and inclusivity that underpin all excellent science.

Over the course of two days, six sessions explored some of the critical emerging themes around the use of AI in science, including its responsible use, the balance between innovation and ethics, a brief overview of the new science it is enabling in a variety of disciplines, and two recent landmark reports on the topic. After each session, Forum participants divided into breakout groups to discuss each topic, with a summary of each discussion then being reported back to the full group. The meeting was not designed to generate conclusions and recommendations, but rather observations and ideas on which future deliberations and actions could build. This document is not a verbatim record, but a summary of the discussions that took place during the event and the key points raised. Comments and recommendations reflect the views and opinions of the speakers and not necessarily those of the Royal Society or the National Academy of Sciences (NAS).

For their work in structuring, organising, and moderating the forum we would like to thank the members of the organising committee, and Anna Bashkirova and Jennifer Clements of the NAS who provided staff support (see Acknowledgments). Videos of the full event are available to view online at royalsociety.org/us-uk-forum-2024

We extend our deepest gratitude to the distinguished speakers, participants, and organisers who made this forum a success. Their contributions have enriched our collective understanding of this pivotal moment in the history of science, and their insights will guide us as we navigate the promises and challenges of AI in the years to come.

Adrian Smith

President of the Royal Society

Marcia McNutt

President of the National Academy of Sciences

Summary

Background

Advancements in artificial intelligence (AI) mark the beginning of a transformative era in science. The increasing use of AI within scientific research enables unprecedented speed and precision within studies, unparalleled approaches to complex problems, and the analysis of vast datasets. It is a catalyst for discovery across disciplines, from genomics and medicine to astrophysics and environmental science.

Yet the integration of AI within science also presents several challenges. Many AI systems operate as ‘black boxes’, where the path taken to reach conclusions is not transparent. Bias within studies may not be readily apparent, and research undertaken using AI models can be difficult to reproduce and replicate. This lack of transparency fundamentally challenges the traditional scientific method that protects and encourages integrity within research. Additionally, the rapid development of AI technology has outpaced considered discussion of privacy, consent to use data and copyright.

These and the many other questions surrounding the use of AI in science were the motivating forces behind the 2024 Forum. Organised by the Royal Society and the National Academy of Sciences, it addressed the emergence of AI in science and its potential for application across scientific disciplines and for developing new scientific methods. It also addressed the concurrent responsibilities that arise when scientists make use of AI, and the practical and ethical challenges of integrating AI within scientific research.

While AI offers tremendous benefits to scientific research, its risk to the established standards of research remains. AI is ultimately a tool, and whether it is useful or detrimental to science depends on its appropriate application. Developers and users of the technology must engage with it responsibly to ensure its benefits are maximised whilst mitigating its risks, enabling a future where technology serves the best interests of humanity and the other species with which we share the planet. This forum was designed to consider how best to ensure this future.

About the National Academy of Sciences

The National Academy of Sciences was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research.

About the Royal Society

The Royal Society is the UK’s national academy of science. It is a self-governing Fellowship of many of the world’s most distinguished scientists drawn from all areas of science, engineering, and medicine. The Society’s fundamental purpose, reflected in its founding Charters of the 1660s, is to recognise, promote, and support excellence in science and to encourage the development of use of science for the benefit of humanity.

The emergence of AI science to support scientific research

This opening session set the scene for the Forum by discussing both the current state of AI as a discipline, as well as how AI can be integrated into scientific research. Jeannette Wing, Executive Vice President for Research and Professor of Computer Science at Columbia University, provided a brief overview of the history of AI, beginning in 1950 when Alan Turing, considered the originator of modern computer science, proposed the Turing test, which is passed when a machine interacts with a human and the latter cannot tell the difference between that interaction and one with another human. The term ‘artificial intelligence’ was then born, as well as the start of AI as an academic pursuit, at the 1956 Dartmouth conference. The first wave of AI that followed was largely based on the symbolic representation of knowledge, and reasoning based on rules, whereas the second wave, which began around the turn of the 21st century, is driven by a much larger amount of data, especially that relating to human behaviour.

Wing went on to argue that the second wave of generative AI modelling has become a tsunami, in which tools based on deep learning can create new (synthetic or artificial) data, in the form of text or images from real-world data. This is having transformative impacts on science, with a key milestone in AI for science passed in 2018 when the generative AI model AlphaFold successfully predicted the structure of proteins, thus solving a long-standing ‘protein-folding’ problem, nearly 60 years after it was first postulated. Other examples cited by speakers in this session include astronomers’ use of AI to detect new exoplanets, predict signatures of new types of gravitational waves, identify a unique object that may be a remnant of two black holes merging, and produce a sharper version of the very first image of a black hole.

Materials scientists are using generative AI modelling to create new material designs, while AI also underpins innovative multidisciplinary collaborations such as the Polymathic AI Initiative, an international team of scientists from physics and astrophysics, mathematics, AI and neuroscience, that is building foundational models – general-purpose AI models trained on vast datasets¹ – that could then be applied, eg, through fine-tuning, to a wide range of scientific problems.

AI has considerable potential in the molecular sciences, particularly in advancing understanding of the electronic structure of quantum chemistry, according to Microsoft Research's Frank Noé. This involves bringing together approaches from machine learning – using neural networks to represent quantum wavefunctions – with traditional techniques in physics such as Monte Carlo methods to develop the 'Deep Quantum Monte Carlo' method, which can solve highly challenging problems involving systems with hundreds of electrons with very high accuracy. The increasing sophistication of AI-based methods such as these, drawing on terabytes of data from large-scale research projects such as Atlas at CERN and the Human Genome Project, are considerable, he argued.

The applications of AI in chemistry were highlighted by the session Chair, Andy Cooper, Professor of Chemistry at the University of Liverpool, including a 'mobile robotic chemist' which can carry out research autonomously, using algorithms to search chemical space for use in techniques such as crystal structure prediction. This is one of many novel approaches in this field of science which brings together human knowledge, AI and robotics technologies to utilise the different strengths of each to address long standing research challenges.

Tom Mitchell, Founders University Professor, Carnegie Mellon University, discussed the research field of human learning, emphasising its longstanding goal of understanding the principles underlying human learning to improve education. He highlighted key findings from behavioural studies, such as spaced practice and active learning, which enhance long-term retention and engagement respectively.

Mitchell explained the potential of AI to revolutionise education by leveraging large-scale online learning environments and intelligent software. He showcased examples from systems like Squirrel AI, which provide detailed log data of student interactions, enabling the modelling of students' knowledge states and predicting their performance on future questions with high accuracy.

He also explored the use of AI in diagnosing students' knowledge gaps and recommending appropriate interventions, such as hints for incorrect answers. By analysing extensive log data, AI systems can identify the most effective hints, improving students' learning outcomes. Mitchell further discussed the role of large language models (LLMs) such as Chat-GPT in enhancing educational strategies. He presented experiments where LLMs were used to predict and generate effective hints, demonstrating their potential to improve upon traditional educational methods.

Pushmeet Kohli, head of AI for Science, DeepMind, discussed the transformative role of AI in advancing scientific discovery. Kohli identified three main trends driving the need for AI and machine learning in science: the scale of experimental data, the complexity of models, and the unique problems AI can address. Kohli detailed the strategic focus of DeepMind's science team, explaining their criteria for problem selection: the significance and difficulty of the problem, the availability of relevant data, and the presence of clear progress metrics. He emphasised a multidisciplinary approach, integrating domain expertise with AI techniques.

Given these examples, Kohli concluded by emphasising the immense potential of AI and machine learning to address contemporary scientific challenges. The case studies he outlined included AlphaFold, for predicting protein structures from amino acid sequences, and AI use in fusion reactor control, where AI has contributed to enabling stable and exotic plasma configurations that were previously unachievable.

He expressed optimism about the future, where these techniques will continue to significantly advance scientific research.

1. Ada Lovelace Institute. *What is a foundation model?* See <https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/> (accessed 27 August 2025).

AI at the frontier of scientific discovery: transformative applications across disciplines

A series of speakers highlighted how AI is transforming the landscape of scientific discovery, as it is rapidly applied across multiple scientific domains, providing new insights and methods while drawing on considerable computational power. They demonstrated a number of these effects as well as the potential of AI technologies to further contribute to scientific knowledge, whilst at the same time highlighting some current limitations.

Tapio Schneider, a climate scientist at the California Institute of Technology, introduced the session by summarising how science has usually progressed over the last four hundred years through a virtuous cycle of iteration, in which the scientist, or scientists, develops a theoretical or computational model, use it to design an experiment or make an observation, which then provides data to learn from to revise or refine the model, with progress possible at any point within this cycle. Increasingly much of this loop can be accelerated and even automated with AI tools, with the potential to automate the entire loop in the future, which was the subject of a recent landmark NAS report on automating research². He highlighted the use of AI in his field of climate modelling, which is using models and current data to predict the increasingly warmer world that will result from greenhouse gas emissions.

Following the emergence of ‘big science’ in the latter half of the twenty-first century, some of the most exciting research being conducted today is ‘mid-scale’, argued Alex Szalay, Director of the Institute for Data Intensive Science at Johns Hopkins University. AI plays a critical role in this, analysing petabytes of data generated by small or medium-sized research groups using advanced technology to develop highly automated robotic experiments.

Kristin Persson, Professor of Engineering at UC Berkeley, demonstrated the importance of making data available to the wider research community to fuel the AI revolution. Many technologies that will be crucial for a sustainable future – such as wind and solar power, energy storage, electric cars and quantum computing, are all reliant on and limited by functional materials. Yet some experimental data in materials science – such as the elastic tensor of some materials – is hard to come by.

The Materials Project at Berkeley stores much of this data in a highly curated form, accessible to the public for free, which has enabled several research groups to use machine learning on the data sets to design new materials. The success of this initiative is illustrated by the fact that every day on average five scientific papers are published which cite the project, remarked Persson.

AI is also helping to bridge the gap between computational chemistry and materials and drug discovery, according to Kim Jelfs, Professor of Computational Materials Chemistry at Imperial College London. One of the biggest challenges in chemistry is discontinuity – ie, a small change in structure having a major impact in terms of a chemical’s properties and/or behaviour. By using machine learning algorithms on large amounts of carefully curated data, such properties can be identified in simulations before being developed in the lab. Such work is also seeing increasing collaboration between chemists and computer scientists, with chemistry PhD students and postdocs increasingly being trained in coding and machine learning both through formal training and through events such as hackathons.

An overview of how machine learning is revolutionising global weather forecasting was provided by Rémi Lam, Staff Research Scientist at Google DeepMind. This is achieved through three main factors, namely speed – particularly in generating predictions; accuracy – through the use of historical weather data, in particular ‘analysis data’, and identifying subtle trends and patterns to predict future conditions; and accessibility – new weather models can be trained and used for generating forecasts without a super computer, and open-sourced models significantly facilitate the emergence of new players in this research field. Machine learning methods can also leverage considerable amounts of untapped data. Lam showcased recent innovations, GraphCast and GenCast, which draw on all of this to predict the Earth’s surface and atmospheric weather in 3D, 10 – 15 days ahead.

2. National Academies (2022). Automated Research Workflows for Accelerated Discovery: Closing the Knowledge Discovery Loop A synopsis was published as *Accelerating Scientific Discovery With AI-Aided Automation*.

Anil Madhavapeddy, Professor of Planetary Computing at the University of Cambridge, highlighted the use of AI in addressing one of the 21st century's most critical challenges, namely monitoring anthropogenic and naturogenic land use in order to better understand and manage the trade-offs between them to balance the needs of humans and nature. It is now possible to provide governments with accessible real-time, natural language interfaces to predict, and assess the opportunity cost of, land use changes. This is enabled by techniques such as quantifying biomass and constructing statistical counterfactuals to measure the causal impacts of human intervention on carbon and biodiversity³, and using real-time query engines underpinned by machine learning to draw on sparse data and update frequently.

Scientific advances powered by AI and new science enabled by AI

A panel of researchers discussed how AI is changing diverse areas of scientific practice, and ongoing challenges that accompany its deployment in scientific research.

David Donoho, Professor of Statistics at Stanford University emphasised how AI is reshaping various scientific fields by speeding up research processes, such as enabling faster data analysis and improving experimental design in disciplines like chemistry and biology. He also pointed out the broader implications of AI in driving innovation, not just by enhancing existing methodologies but also by introducing new ways to solve complex scientific problems.

John Jumper, Director and senior research scientist at Google DeepMind, elaborated on the evolution of AlphaFold, an AI system that accurately predicts protein structures, and the transformative role of AI in biological research. AlphaFold's success builds on decades of curated data from the Protein Data Bank, a crucial resource used for AI model training. Jumper explained how recent innovations in AlphaFold 2 have allowed it to achieve the same accuracy with significantly less data, unlocking the ability to predict previously inaccessible protein structures.

This led to breakthroughs in biological research, such as the Zhang Lab's research at the Broad Institute, where AlphaFold 2 helped solve structural challenges in bacterial protein injection systems, enabling advances in therapeutic research. AlphaFold 3, a collaboration with Isomorphic Labs, has extended capabilities to explore interactions between

proteins, drugs, and DNA, and thus to investigate dynamic and functional aspects of protein structures. Although this advancement faces challenges due to limited data on complex drug interactions, it holds promise for accelerating drug discovery and understanding diseases at the molecular level.

AlphaFold was cited by John Moult, Professor of Cell Biology and Molecular Genetics at the University of Maryland, as a good example of AI's capability in solving complex scientific challenges. Moult co-founded Critical Assessment of Protein Structure Prediction (CASP), a community-driven initiative that provides a standard for evaluating and comparing capabilities and limitations of current methods of modelling protein structure from sequence. Through successive CASP competitions, AlphaFold demonstrated remarkable progress, significantly improving its performance with each iteration. Its success and that of others has enabled the development of highly accurate models that are now being applied in practical contexts, such as diagnosing rare genetic diseases.

Moult also discussed the broader implications of community-driven experimentation, noting successes in protein folding but acknowledging challenges in applying these methods to other complex biological issues, like understanding genetic variants. He is optimistic about the ability of large language models (LLMs), such as ChatGPT, to automate literature retrieval and analysis to support these efforts.

The evolution of AI applications in facial recognition technology was discussed by Jonathon Phillips, Electronic Engineer at the National Institute of Standards and Technology. The presentation traced the field's progress in the past 30 years from the 1993 Eigenfaces approach, a representational learning method in computer vision, to deep learning-based advancements that Phillips demonstrated to have resulted in a significant reduction in error rates from 79% in 1993 to 8% in 2023. Phillips also noted the shift in research focus from verification to large database searches, such as the use of the technology for face recognition by the Federal Bureau of Investigation (FBI). He also addressed the role that open competitions have played in driving innovation and efforts to detect and address racial bias in AI facial recognition. The presentation was underscored by the need to combine human expertise with AI to advance automated face recognition accuracy and the responsible use of the technology.

3. Eyres A et al. (2025). *LIFE: A metric for mapping the impact of land-cover change on global extinctions*. *Phil. Trans. R. Soc. B*. 380 (doi.org/10.1098/rstb.2023.0327).

Caroline Uhler, core institute member of the Broad Institute of Massachusetts Institute of Technology and Harvard, discussed how biomedical sciences can benefit from AI, and likewise how advancements in the field could inspire new foundational approaches to AI. She emphasised that current AI development in her field focuses on prediction accuracy for applications like recommender systems – software tools that provide users with suggestions for items, such as products or cultural content.⁴ However, referencing Judea Pearl’s causal hierarchy model⁵, she argued that addressing complex biological questions, such as uncovering the regulatory logic of cells or anticipating the effects of previously untested interventions, requires AI to move beyond correlation-based prediction and focus on causal interference. Uhler illustrated this with an example from a cancer immunotherapy AI challenge, where algorithms were tasked with predicting how specific genetic perturbations could reprogram T cells. These algorithms were then ranked by their performance, providing a benchmark for evaluating methodological effectiveness and informing future research directions in cellular reprogramming and therapeutic design.

To conclude the session, Andrew Zisserman, Professor of Computer Vision at the University of Oxford and Distinguished Scientist at Google DeepMind, discussed how computer vision can empower scientists by making their data ‘AI-ready’ and automating data annotation. He provided examples across different fields, such as using AI-based face recognition for chimpanzees to study animal behaviour and social networks, analysing X-ray videos of crystal formation in material science to track and detect the crystals in real time, and monitoring penguin populations in harsh environments in Antarctica to track climate change. These applications showcase how AI has unlocked new research opportunities by enabling non-AI scientists to analyse previously inaccessible or unusable data.

Enabling open science, reproducibility, replicability and privacy

Day two of the meeting opened with a session discussing the opportunities and challenges in the conduct of open science in AI-based research, as well as the unique challenges faced in ensuring reproducibility and replicability when using AI-based tools in science.

Victoria Stodden, Associate Professor in the Department of Industrial and Systems Engineering at the University of Southern California, opened the discussion by tracing the evolution of scientific inquiry, from Roger Bacon’s early emphasis on reproducibility and direct experimentation to the current era shaped by AI-driven discovery. She argued that traditional scholarly formats, based on narrative explanation and authorship, are no longer fit for purpose in the age of complex, opaque models.

As the ability to explain discoveries diminishes, the emphasis should shift to the verifiability, precision, and applicability of results, framed within well-defined benchmarks and computationally transparent workflows. Stodden called for a community-led effort to construct an integrated, computable digital scholarly record—one that facilitates reproducibility, supports privacy-preserving data governance, and enables the meaningful evaluation of AI-generated knowledge. Without such a shift, she cautioned, the scientific enterprise risks fragmentation and epistemic drift in the face of accelerating technological complexity.

Sasha Luccioni, AI researcher and climate lead at Hugging Face, followed with a focus on the ethical and environmental considerations of AI reproducibility. She noted that key principles like sustainability and reproducibility are often neglected in favour of performance and novelty. She stated that the lack of reproducibility not only hinders scientific progress but also affects ethical considerations, such as bias detection and fairness improvements.

Luccioni was particularly concerned with the need to evaluate both environmental and societal impacts of AI models. The presentation covered a review of the Machine Learning Reproducibility Challenge and the NeurIPS Code of Ethics, which aims to improve scientific practices and ethical standards. She also discussed Hugging Face’s Energy Star Rating Project for AI model deployment, that aims to guide members of the community to choose models for different tasks based on their efficiency and environmental impact.

4. Science Direct. *Recommender Systems – an overview*. See <https://www.sciencedirect.com/topics/computer-science/recommender-systems> (accessed 27 August 2025).

5. Pearl J, Mackenzie D. (2018). *The book of why: the new science of cause and effect*.

Mark Kelson, Professor of Statistics for Health at the University of Exeter, continued the discussion on reproducibility. Drawing from his experience as interim co-director of the Institute for Data Science and AI, he noted the increasing recognition of widespread non-reproducible research. Kelson proposed adopting systems like those used by the Sainsbury Lab in Cambridge, which involve random checks of preprints and publications within the past three months, to ensure research reliability and adherence to reproducibility standards such as open access to data and code, comprehensive documentation of methodologies, and the use of standardised computational environments.

Kelson also discussed the importance of publishing negative results to provide a complete scientific record. He illustrated this with a student project that tested the reproducibility of papers using open data and found new insights, such as community masking for COVID-19 also reducing flu transmission by about 6%. This example underscored that greater emphasis on reproducibility could uncover valuable findings, benefitting the scientific process beyond verifying initial specific results.

Xiaoxuan Liu, Senior Clinician Scientist at the University of Birmingham and University Hospitals Birmingham NHS Foundation Trust described how, despite the various practical applications of AI in scientific research, it faces a number of implementation and regulatory challenges in healthcare. Liu works across the innovation pipeline – from design to clinical evaluation, regulation and implementation – to explore innovative AI technologies and their clinical effectiveness, focusing on ensuring that AI tools are not only scientifically sound but also safe for patient use. Liu's presentation pointed out some of the difficulties in applying AI innovations in clinical settings, noting that many AI-based devices are not yet widely adopted in Europe and the UK for reasons including lack of regulatory approval. A meta-analysis comparing the diagnostic accuracy of deep learning in medical imaging with that of healthcare professions revealed that only small fraction of these clinical research studies meet high-quality standards required for patient safety and efficacy. This, she argued, contrasts with the proliferation of often unregulated direct-to-consumer health apps and the media's tendency to emphasise AI's potential benefits in medical care while overlooking its limitations.

To address these issues, Liu's team contributed to the CONSORT-AI (Consolidated Standards of Reporting Trials–Artificial Intelligence) extension, a reporting guideline for clinical trials with an AI component. This guideline aims to promote transparency and completeness in understanding, interpreting and critically appraising clinical trials involving AI interventions. Liu also discussed the STANDING Together (STANdards for data, to ensure Diversity, Inclusivity and Generalisability of AI in health) initiative that seeks to address bias in datasets used to build AI models by establishing standards for data documentation and evaluation. This project has identified significant gaps in documenting critical data characteristics such as race and ethnicity, with less than 5% of skin cancer datasets reporting skin colour. These efforts have informed regulatory practices, contributing to updated standards at the UK's Medicines and Healthcare products Regulatory Agency (MHRA).

Rebecca Lawrence, Managing Director of F1000, concluded the session by exploring how AI can support the integrity of the scientific record, addressing the growing challenge of verifying research output accuracy. AI tools are increasingly being used by authors and peer reviewers alike. In the case of the former, there is a risk of errors – such as biased or incomplete data – influencing published findings. However, AI also offers tools for improving content verification, detecting image manipulation, and enhancing language equity for non-native speakers. Lawrence suggested that by using AI-driven real-time verification systems, researchers can better ensure the reliability of their work while maintaining transparency and reproducibility in academic publishing. However, in her view, it remains important that human-led decisions should guide the use of AI, particularly in identifying concepts, verifying statistical methods, and consulting experts when necessary.

Panel on recent publications

This session examined two pivotal reports on the impact of artificial intelligence (AI) on scientific research. The discussion focused on the President's Council of Advisors on Science and Technology (PCAST) report, *Supercharging research: Harnessing artificial intelligence to meet global challenges*, and the Royal Society report, *Science in the age of AI: How artificial intelligence is changing the nature and method of scientific research*. Both reports offer critical insights into how AI is reshaping research practices and methodologies.

The President's Council of Advisors on Science and Technology (PCAST) report, *Supercharging Research: Harnessing Artificial Intelligence to Meet Global Challenges*, was introduced by Professor Maria Zuber, Professor of Geophysics at Massachusetts Institute of Technology and PCAST co-chair. PCAST, an advisory body to the President of the United States, was tasked with providing strategic insights in response to President Biden's Executive Order on the Safe, Secure, Trustworthy Development and Use of Artificial Intelligence. In this context, the report focuses on how AI can support scientific research, its role in addressing global challenges, and the safeguards needed to ensure its responsible use.

Zuber reflected on the methods used to compile this report, stating that the report emerged from extensive consultations with experts in AI, scientific research, and policy, and takes a broad approach to evaluating AI's potential. It recognises both the positive opportunities AI presents in accelerating scientific discovery and the challenges it poses, particularly regarding ethical use, data privacy, and reproducibility. While its recommendations are directed primarily toward the President and federal agencies, the findings are highly relevant to other stakeholders in academia, industry, and civil society.

Professor Terence Tao, Professor of Mathematics at the University of California and PCAST co-lead then addressed the key themes of the report. The first theme, human empowerment, emphasises that AI should complement, not replace, human scientists. The report highlights that AI's potential lies in augmenting human capabilities by automating routine, labour-intensive tasks such as data processing, allowing researchers to focus on more strategic and high-level decision-making. This approach aims to 'supercharge' research productivity, with AI functioning as a tool to accelerate breakthroughs in fields like materials science, therapeutic drug discovery, and semiconductor design.

Another central theme highlighted by Tao is the responsible use of AI, which stresses the importance of transparency, replicability, and environmental sustainability in AI-driven research. The report warns of the risks associated with AI, such as bias, opacity in decision-making, and environmental costs from high computational demands. It calls for implementing responsible AI practices, including ensuring that AI results are replicable and open to external validation. Additionally, the report advocates for practices that reduce AI's environmental footprint, such as optimising energy use in computational models and ensuring AI infrastructure is sustainable.

The third key theme Tao mentioned is shared and open resources, which focuses on the necessity of developing and sharing open-source AI tools and infrastructure. The report recommends expanding access to AI technologies and datasets, particularly through initiatives like the National Artificial Intelligence Research Resource (NAIRR), a pilot project led by the National Science Foundation (NSF). By making federal datasets more accessible and improving data-sharing practices, the report aims to democratise AI capabilities, preventing monopolisation by a few powerful entities. Tao emphasised that this would enable broader collaboration across scientific fields and between industry, academia, and government institutions, ensuring that the benefits of AI are widely distributed.

Professor Alison Noble, Technikos Professor of Biomedical Engineering at the University of Oxford, Foreign Secretary of the Royal Society and chair of the *Science in the age of AI* report working group, presented the Royal Society's report, *Science in the age of AI*, that provides an examination of how AI is altering the nature and methodology of scientific research. The report is the result of extensive consultations with over a hundred leading scientists from the Royal Society's Fellowship and other network members. These insights were gathered through interviews, roundtable discussions, and global workshops, including collaboration with the Chinese Academy of Sciences and participation in AI safety meetings. Noble noted that the report was a response to the growing influence of AI and identified both opportunities and challenges that this new technological era presents for the scientific community.

Noble outlined key themes from the report, beginning with the evolving skills landscape, where researchers increasingly prioritise AI-related competencies, often at the expense of traditional scientific expertise. This trend, while important for leveraging AI's potential, raises concerns about maintaining a balance between new technological competencies and core scientific understanding. Much like earlier advances in technology, this shift is challenging the structure of scientific training. Noble suggested that the report calls for new recognition systems for emerging roles like data curators and information managers, who are becoming increasingly crucial in AI-driven research environments.

Another theme in the report is the increasing need for interdisciplinary collaboration. AI's impact spans multiple scientific fields, requiring researchers to break traditional disciplinary boundaries and work more closely with AI specialists. Noble pointed out that interdisciplinary teamwork is essential to fully unlock the benefits of AI in research, echoing the growing call for collaboration across different scientific sectors.

When discussing areas of action, Noble identified three critical aspects: access, trust, and the ethical use of AI. The report stresses the need for equitable access to AI tools and infrastructure, recognising that uneven availability and access to this technology hinders its integration into research where needed. Enhancing educational curriculums to improve AI literacy and the usability of the tools is also crucial. The second area highlighted was trust in AI-generated research outputs, with the report calling for transparency, reproducibility and robust verification methods to ensure reliable results and thereby greater public trust in science. Finally, ethical AI uses were highlighted as essential, particularly in sensitive areas like healthcare, where the development of ethical guidelines is necessary to address bias and maintain fairness and transparency.

The session speakers concluded that both reports underscore the need for a balanced approach to integrating AI into scientific research. They emphasise the importance of maintaining scientific integrity, ensuring the responsible and ethical use of AI, and improving access to resources. The subsequent discussions reflected a shared commitment to leveraging AI's potential while addressing its challenges to ensure that research remains robust, reliable, and beneficial to society.

Responsible AI in science

This final panel session saw speakers offer their perspectives on the measures that should be taken to ensure AI in science is used responsibly and ethically, in ways that support the use of science to benefit humanity.

Yolanda Gil, Research Professor of Computer Science and Spatial Sciences at the University of South California, discussed the potential for further integrating AI scientists as partners in scientific discovery. She questioned the current role of AI in the scientific ecosystem and highlighted the potential benefits of more integration. Gil identified several shortcomings in human-led scientific research, including a lack of systematic approaches, errors, biases, and poor reporting. She used the example of a study on fossil records to highlight how some tasks can be undertaken more effectively by AI than humans. She also presented examples such as the retraction of papers due to errors and the limitations of human bias in hypothesis generation to demonstrate further strengths of incorporating AI into scientific research. When considering the future of AI in science, she envisioned that AI systems could write scientific papers, help reproduce published articles, and act as collaborators in research. She suggested that AI could be used to automate complex workflows and improve reproducibility and open science principles.

Stu Feldman, President and Chief Scientist at Schmidt Sciences, addressed the challenges and impacts of AI on scientific norms, openness, and the pace of research. He acknowledged the transformative effects of AI technologies, such as neural networks and language models, on various scientific fields and the need for serious consideration of their implications.

Feldman emphasised the radical changes AI has brought to fields such as computer vision, where traditional methods have been replaced by AI-driven approaches. He noted that AI can exacerbate previous concerns around scientific research (bias, reproducibility, pressure to publish, etc), necessitating a deeper understanding of why failures occur when using AI and how to address them. He shared insights into Schmidt Sciences' funding of postdoctoral researchers using AI to support non-computer science research, illustrating the growing interdisciplinary nature of AI applications. He also raised concerns about the reproducibility of AI research, given the inherent randomness in AI models, and the challenges of aligning the rapid pace of AI advancements with traditional scientific methodologies.

David Leslie, Director of Ethics and Responsible Innovation Research at the Alan Turing Institute, discussed some of the myths and realities of Artificial General Intelligence (AGI), attempting to dispel some of the exaggerated expectations and myths surrounding AGI (such as whether it ‘thinks’), and emphasising the importance of maintaining a realistic perspective on AI capabilities. He highlighted the ethical challenges posed by AI, stressing the need for responsible development and deployment. Leslie also called for continuous public engagement and transparent disclosure in AI research in order to build trust and ensure ethical standards are upheld, underscoring the importance of integrating ethical considerations into AI research and development, and advocating for a balance between technological advancement and societal impact.

Shannon Vallor, Baillie Gifford Chair in the Ethics of Data and Artificial Intelligence at the Edinburgh Futures Institute, discussed the need to integrate ethical considerations into AI research. She argued for broader metrics to evaluate research impact and called for collaboration with policymakers and public sector funders to protect research integrity and prevent undue influence from commercial interests. In particular, Vallor advocated for the inclusion of sustainability, ethics, and rigour in research evaluation metrics, alongside traditional impact measures. In calling for a holistic approach to AI research, she emphasised the potential role of the humanities in meeting the ethical challenges of AI development.

During the Q&A, Vallor commented on the hallucinatory tendency of generative AI. She cautioned that generative AI is not always the most appropriate tool for every research task. She also highlighted that generative AI is unable to produce anything truly novel, and that its reliance on the most dominant data pattern affects its output. Many primary developers of AI technology are private actors with commercial interests, and governments are already interested in the potential for AI in surveillance, among other things. She expressed concern over the extent to which public trust and understanding of AI could be built under these circumstances.

The final speaker was Baron Rees of Ludlow, Astronomer Royal, who discussed the applications and challenges of AI in astronomy and fundamental physics. Rees emphasised the transformative impact of AI on handling vast amounts of astronomical data, which previously were unmanageable for researchers. He used the example of autonomous robots in space exploration to underscore the potential for AI to enable more cost-effective scientific missions compared to manned missions. Rees ended by speculating on AI’s role in theoretical advancements, particularly in scenarios where human intuition may fall short. AI could potentially uncover new theories or validate existing ones by exploring vast mathematical spaces beyond human capability, including theories relating to the origin of the universe.

Acknowledgements

The members of the organising committee are listed below. Members contributed to the Forum on the basis of their own expertise and good judgement.

Forum organising committee

Alison Noble, Foreign Secretary and Vice-President, the Royal Society (co-chair)

William Press, Treasurer, National Academy of Sciences (co-chair)

Andy Cooper, Professor of Chemistry, University of Liverpool

David Donoho, Professor of Statistics, Stanford University

Chris Holmes, Professor of Biostatistics, University of Oxford

Francesca Rossi, IBM AI Ethics Global Leader and Distinguished Research Staff Member, IBM T J Watson Research Lab

Tapio Schneider, Theodore Y Wu Professor of Environmental Science and Engineering, California Institute of Technology

Reviewers

This document was reviewed in draft form by the co-chairs. In addition, Jeannette Wing, Executive Vice President for Research, Columbia University, served as arbiter to ensure that review comments were appropriately incorporated into the final version of this summary. The Royal Society and National Academy of Sciences gratefully acknowledge the contributions of the reviewers.

Royal Society and NAS staff

Many staff at the Royal Society and NAS contributed to the forum and the production of this summary.

The project team are listed below.

Royal Society and NAS staff

Areeq Chowdhury, Head of Policy – Data, the Royal Society

Luke Clarke, Head of International Affairs – Americas, International Organisations and Africa, the Royal Society

Poppy Joyce, Policy Adviser – Americas and Academic Freedom, the Royal Society

Nicole Mwananshiku, Policy Adviser – Data and Digital Technologies, the Royal Society

Ashleigh Carver, Manager, Scientific Meetings, the Royal Society

Maria Fernandes, Scientific Programmes Officer, the Royal Society

Anna Bashkirova, Director, Special Programs, National Academy of Sciences

Jennifer Clements, Program Coordinator, National Academy of Sciences