Critical Issues for the Uptake of AI in Science

The following list of critical issues is a result of a review of over 300 publications on the integration of AI in national research ecosystems. The list is an extract from the working paper 'Preparing National Research Ecosystems for AI: strategies and progress in 2024', https://council.science/publications/ai-science-systems.

THEME 1: R&D agenda setting, technology assessment, foresight and science advice

› PRIORITY SECTORS

a. Priority-setting

- We must find ways to identify strategic sectors for AI development and for its uptake by the scientific community. Mechanisms may include funding, infrastructure development and capacity building programmes.

› FUNDING PRACTICES

a. Will AI capacity replace scientific merit in science funding decisions?

- AI intensity may become an inappropriate deciding factor in determining the allocation of resources and hence the trajectory of scientific discovery. Its salience could close off areas of research that do not use it.
- Competition within research could become less a matter of merit and more a matter of access to AI. This risks poor decision-making and further concentration of research funding.

b. Use of AI in resource allocation

- AI relies on machine learning from existing material. It may produce reviews that are inherently conservative and which reproduce old biases.

c. Impact of AI on evaluation panels

- AI-driven science tends to be interdisciplinary because AIs do not know subject boundaries. Today's domain-led expert panels may be unable to review it adequately, despite the many recent calls for science to be more interdisciplinary.

› CAPACITY BUILDING AND RETENTION

a. Growing AI skills in the scientific community

- There is a need for broad but differentiated AI skills development for learners and practitioners at all levels. Important aspects include education in AI, training in domain-specific use, ethics, and interdisciplinary competencies. Teaching will have to recognise that this is a fast-moving topic.

b. Diversity in AI research

- There is a need to ensure the gender, ethnic and cultural diversity of the AI workforce, in the interest of equity and to improve the quality of research and other outcomes. Machine learning can reproduce existing inequity.
- We have to develop the right incentives for disciplinary and interdisciplinary AI.

c. Talent retention in the public science sector

- Public sector science, including universities and research centres, needs talent acquisition and retention, given the strong demand for AI skills from the private sector. Unusually, this is an area in which the private sector can offer interesting jobs as well as high salaries.

› INFRASTRUCTURE

a. Development of cloud computing appropriate for science

- Uncertain funding for cloud computing and research data repositories constrains scientific advances. In the absence of public cloud capacity, wealthier research institutions are likely to contract private companies, limiting the sharing of their research data and leaving less wealthy institutions behind.

b. The digital divide goes algorithmic

- We must determine how inequity in AI access between individuals, groups, academic disciplines, organisations and locations results in poorer research outcomes.

c. Development of AI tools for science

- We must determine what kinds of partnerships will encourage the development of AI tools appropriate for specialized research institutions. How do we ensure that new AI technologies are not driven solely by the AI and machine-learning communities, but rather developed jointly with all research communities?

› INTERNATIONAL COLLABORATION

a. Variation between legal systems

- We need to assess how jurisdictional variability in governance and data protection between countries impacts international research and research collaboration.

b. Regional collaboration

- Countries must find out the extent to which they can cooperate to establish regional AI centres and research networks if they don't have the resources to do it on their own.

› JOBS, CAREERS AND EMPLOYMENT

a. Impact on jobs in science and research

- There is a need to monitor how advances in AI affect the number and nature of jobs in science.

b. Continuous AI training

- There is a need to develop ways for scientists and research staff to keep up to date with AI in order to produce better research and minimise job losses. There may need to be specialist AI trainers and teachers, for example to help users understand the ethical issues raised by AI.

› NETWORK AND REPOSITORY SECURITY

a. AI effects on scientific cybersecurity

- Science institutions must ensure the best possible network hygiene, ensure the security of partner organisations, and control cybersecurity risks from individual people. How do they secure facilities against intellectual property theft, access to private and sensitive data, and ransom attacks?
- The protection of data quality and integrity requires controls on access to repositories, as well as highly qualified personnel, strong partnerships and an appropriate built environment.

THEME 2: Public engagement, science communication and public accountability

› SCIENTIFIC INTEGRITY IN THE CONDUCT OF RESEARCH

a. Principles and values of current science

- AI may generate tensions between some of the core principles and values that define today's science. Such contradictions might include openness vs. rigour; privacy and confidentiality vs. open science; massive data vs. high quality data; or explainability vs. "black box" results.

b. Reliability and explainability of results

- Lack of trust in AI, within science and in other activities, may create challenges for its uptake in science. But uncritical trust will lead to a potentially dangerous overreliance on AI technology and the results it generates. AI tends to produce normative results rather than groundbreaking insights, because it is based in existing knowledge and existing opinion.

c. Reproducibility

- Today's science already has severe reproducibility issues. How will AI worsen them or perhaps solve them? For AI to improve reproducibility it will need to be more transparent, providing more information about codes, underlying data and experiment design. This applies both to AI research and to research using AI.

d. Explainability of results

- The scientific method requires scientific claims to be explainable and understandable. Some popular AI methods operate as a black box, making it impossible to say how they have reached their conclusions or to identify spurious correlations or causalities.

e. Ethical data use

- The use of big data and AI complicates present-day notions of consent and of human research participants, as well as the ways in which data is collected and used.
- AI Ethics and Review Boards focus on human subjects. As well as carrying out their present vital role, they should be able to examine possible harms to wider society.

f. Accountability

- We will have to determine who is responsible for fabrication, falsification, plagiarism and other bad practice when the faulty conduct can be traced back to an AI. The answer may be simple if the AI has an obvious owner, but in the future many may not.

g. Conflict of interest

- We need to see whether new conflicts of interest arise as AI spreads. They may not be covered by current conflict-of-interest policies.

› ENVIRONMENTAL IMPACT

- AI development has to be made more sustainable (in relation to the use of computer chips and electricity in particular). More fundamentally, AIs may well not be attuned to environmental concerns if they have not learned from appropriate input materials.

› SCIENTIFIC PUBLISHING

a. Acknowledgment of contributors and authors

- Researchers have to explain how AI was used in the production of research outputs.

b. AI for policing science

- Publishers have to determine whether AI should be used to detect non-AI generated fabrication, falsification and plagiarism.

THEME 3: Regulation, standards, private sector governance and self-regulation

› DATA QUALITY

a. Accuracy

- Larger datasets are better for training AIs, yet they are also more likely to produce responses based too closely on the data available to them (overfitting) or to contain inaccuracies and biases that could result in wrong or misleading results. Incorrectly sourced data, Frankenstein datasets and biased datasets already have dangerous implications for science. This problem needs to be addressed at every level, from considerations of governance and management to operational use.

b. Bias and exclusion

- While AI, and large language models in particular, use 'biases' (statistical similarity) in data to produce results, it is important to curate training data to avoid further marginalization of particular groups and regions. Digital exclusion leads to gaps in data. Furthermore, how do we represent those who are offline?

c. Subject orientation of data vs. the interdisciplinary nature of AI research

- Most scientific knowledge comes from a specific subject. We need to encode and use it, while enabling communication between domains and allowing for the growing generation of interdisciplinary knowledge.

d. Data coding and annotation

- AIs, and large language models in particular, require humans to code and annotate the data they use. These individuals must be aware of the risk of embedding cultural differences in the data during the annotation process.

› DATA MANAGEMENT AND GOVERNANCE

a. Open data vs. AI safety

- Access to high-quality data is crucial to the development of AI for science. But the public interest, as well as that of individuals, calls for governance structures to protect privacy and to guarantee the ethical use of data.

b. Access vs. Advantage

- Much of the data required for the development of scientific AI will not fall within the scope of open data initiatives, for example data held by the private sector. The tension between enabling access and maintaining commercial advantage may result in high-quality data being kept confidential.

c. Data infrastructures

- The development of AI for science will require harmonization of practices and the development of communities of practice. Current norms and practices for the

- production and use of data differ between disciplines and institutions.
- As scientific organizations increase their data curation and storage capacity, they will need to increase interoperability between repositories.

› DATA STANDARDS

a. Data standards for provenance

- The sources of training data must be appropriately disclosed and evaluated. A specific concern is the ethical aspect of data and data sources, and its implications for bias in AI.

b. Data standards for quality (see also 'data quality' above)

- Technical standards, certification and compliance should be imposed to ensure that data used in science is properly curated and stored.

› LAW, REGULATION AND POLICY

a. Legal liability of research done with AI

- We have to reconcile traditional liability systems with AI processes and outputs, with their varying degrees of autonomy and transparency. At what point does an AI, rather than its maker, become responsible for its actions?

b. Copyright protection or patenting for machine-generated creations?

- Uncertainty about the eligibility and appropriateness of copyright protection for AI-generated creations may lead to the use of patenting or trade secrecy techniques to protect intellectual property. This would reduce public availability of the valuable results, positive and negative, of AI projects.

c. Protection and use of digital data

- Text and data mining risk infringing copyright through the creation of unauthorized copies, and may violate the terms and conditions of websites and databases. The United Kingdom is creating a copyright exception rule for text and data mining, and other jurisdictions may follow.
- Works mined for data can be protected by copyright, but data themselves are usually protected only if they were part of original datasets. This may lead to the use of trade secret to protect data. The European Union protects data extracted from protected databases for scientific research. But the borderless character of digital data exacerbates tensions between jurisdictions.

› REGULATIONS

a. The domestic regulatory environment

- Work towards domestic AI regulation will be a balancing act between different considerations and needs. In these arbitrations, countries must create beneficial conditions for their science and research sectors to thrive and work for the common good.

b. Impact of regulation in other jurisdictions

- Observation of other countries' actions can lead to leap-frogging and the alignment of provisions; or, uncertainty about regulation may lead some legal regimes to seek competitive advantage through less rigorous regulation, to the detriment of the country where the creation was generated.