Preprint:
Newman-Griffis, D. orcid.org/0000-0002-0473-4226 (Submitted: 2023) AI thinking and the enterprise of science. [Preprint - OSF Preprints] (Submitted)

https://doi.org/10.31219/osf.io/wb83e

AI thinking and the enterprise of science

Denis Newman-Griffis
Information School, University of Sheffield, Sheffield, UK
Research on Research Institute, London, UK

Abstract

Rapid developments in artificial intelligence and the mainstreaming of generative AI have raised vital questions about the future of science. AI techniques present significant potential for enhancing scientific research, but also risks of bias, inequity, and eroding originality. Approaching AI as a methodology rather than a technology can help manage the tension between promise and peril, anchoring AI efficacy and ethics in processes of AI design and use. An "AI thinking" perspective can help scientists achieve richer analysis, more open science, and constructive use of generative AI while helping manage risk of harm. Adopting AI thinking requires diverse efforts in education, AI tools, and new public narratives, but these efforts will be rewarded with new approaches to AI as a fundamental toolbox for contemporary science.

Introduction

We are currently faced with the challenge of redefining what science looks like in an era of commonplace AI. As a general strategy for using knowledge and data to help perform tasks with computers, the umbrella of "AI" includes a wide range of technologies and applications and has come in many forms over the decades since its inception. Most recently, Generative AI—with its capacity to generate text, images, and other forms of data—is posing new challenges and opportunities in science (*1*), adding to long-standing questions about the ethics and management of AI and machine learning in science (*2*).

In the current AI landscape, the key challenge to the scientific community is to map out: How do we use AI effectively in science, while managing its risks and understanding its limitations? How is the increased use of AI in science likely to impact science policy makers and end users of scientific research? And what shape does the enterprise of science take in a more AI-powered world?

The greatest potential for AI as a force for innovation in science, and the most powerful way to manage its use and misuse, lies in recognising AI as a *methodology* for working with data and tasks, rather than a set of technologies and tools alone. AI has the potential to be for twenty-first century science what statistics was for science in the nineteenth and twentieth centuries: like the universal toolbox that statistical methods provide for working with evidence in research (*3*), AI thinking can be a methodological toolbox for working with data and information in research. And just as statistical thinking can help measure how scientific evidence represents and impacts

Correspondence to: Denis Newman-Griffis, d.r.newman-griffis@sheffield.ac.uk

broader society, so AI thinking can help put data and its use into a broader picture of ethical and social impact.

What is "AI thinking?"

AI thinking reflects the *process* of how we design and use AI systems to solve problems in specific contexts. AI systems are technologies that use knowledge about the world to process and analyse data in a human-like way. This may involve hard-coded expert systems as well as data derived from observation via machine learning. AI in practice often uses both, such as in chatbots or automatic speech recognition.

Specific systems and software that implement these kinds of processes can be thought of as "AI tools." A diagnostic model for cancer imaging, a generative text engine like ChatGPT, an optimisation model for shipping logistics—each of these is an AI tool, built to do a particular thing at a particular point in time. AI tools operate on an input-output principle: given a particular kind of input (a radiology scan, a text prompt, a set of orders), they are built to produce a useful output (a diagnostic prediction, a text response, a routing). These input-output relationships can be tremendously powerful, and the last two decades have seen astounding performance of specific tasks by AI tools. However, these tools are unavoidably limited by the data they work with, the perspectives they reflect, and the input-output relationships they have been exposed to, and each of these risks harm when left unacknowledged or unaddressed (*4*).

AI thinking describes the ways we put AI tools in context. It reflects the process of breaking down a potential use case for AI into smaller pieces to define what you want to accomplish, what information you have to inform that goal, and how the use of AI informs and affects the people and processes you are working with. Table 1 illustrates aspects of the AI thinking process for three example use cases of AI in science: in scientific discovery, data processing, and authoring scientific manuscripts. Taking an AI thinking approach enables us to map out what practical AI use might look like for these goals and how we can ground use of AI in scientific understanding and process. This shows how use of AI is only one part of a bigger picture, and illustrates how we can put AI tools into goal-driven—rather than technology-driven—contexts.

Why use AI thinking in science?

Richer analysis

Computers excel at processing large volumes of data. AI systems take this further by enabling knowledge-driven approaches to analysing and combining data that are highly *complex* as well as high-volume: merging information from multiple sources that capture different facets of complex phenomena, such as combining medical imaging with health records and lab readings (*5*). Statistical modelling also offers tools to help manage issues of uncertainty and limited representation across data sources (e.g., due to data provenance), though these must also be carefully considered in the design of AI systems to use such data.

Table 1. Illustration of using AI thinking to map out the information, process, and context of using AI for three example use cases in scientific research. The examples are neither exhaustive nor prescriptive, but serve to illustrate the range of questions and decisions involved in situating AI use for specific purposes in science.

Part of science enterprise	**Knowledge discovery**	**Data processing**	**Research outputs**
Example goal	Identify potential cancer precursors from large-scale biomedical data	De-noise astronomic measurements and merge images	Generate informative captions for tables and figures in scientific articles
Sub-problems	• Find groups of biomedical measurements in large populations • Model relationship between measurement groups and cancer rates	• Detect measurement outliers and/or errors • Generate composite images	• Use figure/table content to generate text description • Measure accuracy and informativeness of generated captions
Relevant information	• Genetic features • Health records • Biological data • Lifestyle features • Social determinants of health	• Source measurements/images • Known distributions and parameters for measurements • Measurement sources/types	• Table headers • Table rows • Figure image • References in manuscript text
Information context: key questions	• What groups are or are not represented in the data? • What differences are there in access to care or measurement that might affect the data?	• Are the "known" measurement parameters applicable to these samples?	• What language is the article written in? • What audience is the article written for? • How do these match up with off-the-shelf generative AI tools?
How to use AI to draw on information	• Cluster biomedical data based on similar patterns • Develop predictive models to estimate cancer risk given input biomedical measurements	• Use known parameters to estimate likelihood that sample measurements are valid, outliers, or errors • Match portions of multiple images together to create composite image	• Use ChatGPT or other generative AI to process the figure or table and produce a draft caption • Use translation tools to translate table content and figure captions between languages
Context of AI use: key questions	• Who is represented in the implementation and use of the AI? • How are AI outputs and predictions validated and verified? • How will this inform decisions about cancer research or care?	• Will classification errors mean important findings are missed by being labelled as measurement errors? • Are AI-processed data perceived as valid or trustworthy by peers?	• How to ensure captions are manually reviewed for accuracy? • How does this interact with accessibility goals and concerns? • What are journal policies on originality in re generative AI use?

The process-oriented model of AI thinking enables scientists to view this capability for analysis from a situated, context-driven perspective, and to map questions of data provenance and trustworthiness to specific design decisions. Key questions to guide this process include: What information do I think might be relevant to my research? What can I gain scientifically from integrating deeper and more diverse sources of information, such as measurements of the same phenomenon taken by different researchers with different methods? AI tools make it possible to analyse deeper and richer combinations of data, while the questions that drive that data selection and analysis remain grounded in expert understanding of the science.

More cumulative open science

AI thinking can also open new avenues for a fundamental challenge of science: building cumulatively on the work of peers and predecessors. The growing adoption of open science principles is making more scientific literature and research data available than ever before (*6*). However, the challenges of standardising data infrastructure and a lack of appropriate methodologies to handle the conceptual and measurement differences of heterogeneous research data have limited the impact of open data as a driver for scientific advances (*7*).

AI methodologies can help to combine diverse sources of open research data—including in different formats—and integrate them into unified wholes, through techniques such as the abstraction and interrelatedness provided by deep learning. AI methods can also be used to leverage open research literature at super-human scale to understand patterns not previously accessible in science (*8*). A knowledge-driven AI thinking approach provides the framework to put these processes in context, ensuring that the selection and integration of data and interpretation of findings is grounded in appropriate scientific understanding. Through combining technical analysis with situated understanding, AI thinking can help to realise the potential of open science to drive richer and more cumulative learning.

Constructive generation

AI use is never a straightforward good or ill, and this is rarely clearer than with generative AI. Generative AI technologies make it much easier to fabricate plausible data, images, and text; they pose substantive risks for scientific integrity, and it is vital to develop standards for ethical use of generative AI in research outputs. At the same time, these technologies have enormous potential as assistive tools for science: e.g. for initial background information to start exploring a new topic, improving the ease and readability of writing scientific prose—or translating it between languages—even generating potential counter arguments or framings, serving as a sort of statistical constructive critic. The challenge for the scientific community is therefore to maintain quality and integrity while embracing emerging possibilities for new ways of pursuing our scientific goals (*9*).

The situated perspective of AI thinking is key to productively managing this dilemma. When we view large language models as statistical extrapolation of what a large sample of human language might "say" to a given question, and view this capability through specific knowledge of disciplinary norms and practices, the risks become more manageable and the possibilities more achievable. This way of thinking can also broaden beyond text to help with exploring opportunities and challenges in image generation and multimodal AI.

Responsible AI frameworks

The focus on process and decisions about design and use that is the heart of AI thinking also provides a basis for critically analysing how AI is used and misused in practice, and understanding where and how ethical risks arise. As AI is used more and more widely, it only becomes more urgent to develop robust ethical and responsible practices around using and managing AI technologies. Auditing methods, explainable AI, bias mitigation and other approaches are valuable methodological tools, but too often treat ethical and responsible AI as a technological problem in need of a technological fix (*10*).

An AI thinking perspective reframes the question in terms of ethical and responsible decisions in the *design and use* of AI. This places assessment of AI risks, as well as AI governance, in a proactive focus on the social contexts where AI systems are designed and used, rather than a reactive, technology-centred perspective. This contextualised perspective anchors the questions of where ethical risks and failures arise and what we can do about them in specific decisions made by people and organisations *during* AI design and implementation, instead of *after* systems are extant and released into the world (*11*). AI thinking can therefore make management of AI risk more tractable and actionable both for internal management and external governance.

What do we need to enable adopting AI thinking?

Educational innovation

Current AI education emphasises teaching learners to build, evaluate, and improve AI technologies. Supporting broader use of AI with AI thinking perspectives requires new educational approaches focusing on how to work with AI in context and understanding its offerings and limitations. Future-oriented AI education must also speak to the needs of diverse audiences who engage with AI in non-technical ways. This will include the core ideas of AI thinking and how AI use affects daily life, as well as the skills and practices of AI thinking and understanding the use of AI in context. It is also vital that principles of AI governance, ethics, and decision making are built into the foundations of contemporary AI education, and an AI thinking approach provides the tools to anchor these principles in teachable practice.

Usable AI tools

For AI thinking to be effective, AI methodologies must be accessible to non-experts. Recent development of machine learning toolkits has made complex AI methods much more accessible and contributed significantly to the current AI boom (*12*). But using these toolkits still requires significant expertise in programming and large computational resources, as well as specific paradigms for data and analysis, excluding many scientists and much of the Global South (*13*). ChatGPT illustrates the transformative power of simple, intuitive interfaces for AI technologies, though at many costs in technology transparency and control. For AI to become a generalpurpose toolbox that is more accessible across expertise and access to computational resources, new software packages, visualisation interfaces, and more natural AI workflows are needed.

New public narratives

Public narratives about AI have changed significantly over the last decade, most recently highlighting existential fears. However, most AI narratives focus on the AI technologies and little on the roles and responsibilities of the people, processes, and perspectives designing and using them (*14*). Building narratives around an AI thinking perspective can help shift the focus from AI technologies to how they are created and used. This reframing can help demystify the risks and overwhelmingness of technology-driven narratives and highlight ways that everyone—scientists and engineers, policymakers, and the public—can understand, work with, and critique the use of AI as a tool for scientific research.

Conclusion

Just as statistical thinking transformed how scientists defined and worked with evidence in nineteenth and twentieth-century science (*15*), so AI thinking can help transform how scientists work with data in the twenty-first century. Science always requires 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.

Realising the benefits of AI in science requires an AI thinking approach, grounded in the contexts and processes in which AI systems are designed and used. Adopting an AI thinking approach makes AI technologies more tangible and tractable to use in practice, and also provides invaluable anchor points for understanding the factors making AI use both effective and ethical. Achieving a shift to AI thinking needs concerted efforts in education and training, usable AI tools, and reframing AI narratives, but these efforts will be rewarded with a richer, more flexible, and more interlinked scientific enterprise of the future.

Acknowledgments

We gratefully acknowledge Xin Zhao, Andrew M. Cox, Michael Thelwall, Stephen Pinfield, the organisers and attendees of WOEPSR 2023, and the participants and partners of the Research on Research Institute GRAIL project for invaluable discussions and feedback on the arguments presented here.

References

1. A. Birhane, A. Kasirzadeh, D. Leslie D, S. Wachter, Science in the age of large language models. *Nature Reviews Physics* **5** (5), 277-60 (2023).
2. M. I. Jordan, T. M. Mitchell, Machine learning: Trends, perspectives, and prospects. *Science* **349** (6245), 255-60 (2015).
3. D. R. Cox, B. Efron, Statistical thinking for 21st century scientists. *Science Advances* **3** (6), e1700768 (2017).
4. E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery, 2021), 610-623.
5. A. Barua, M. U. Ahmed, S. Begum, A Systematic Literature Review on Multimodal Machine Learning: Applications, Challenges, Gaps and Future Directions. *IEEE Access* **11**, 14804-14831 (2023).
6. V. Gewin, Data sharing: An open mind on open data. *Nature* **529** (7584), 117-119 (2016).
7. L. Himanen, A. Geurts, A. S. Foster, P. Rinke, Data-Driven Materials Science: Status, Challenges, and Perspectives. *Advanced Science* **6** (21), 1900808 (2019).
8. A. P. K. Nelson, R. J. Gray, J. K. Ruffle, H. C. Watkins, D. Herron, N. Sorros, D. Mikhailov, M. J. Cardoso, S. Ourselin, N. McNally, B. Williams, G. E. Rees, P. Nachev, Deep forecasting of translational impact in medical research. *Patterns* **3**, 100483 (2022).
9. C. Stokel-Walker, R. Van Noorden, What ChatGPT and generative AI mean for science. *Nature* **614** (7947), 214-216 (2023).
10. L. Sartori, A. Theodorou, A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics and Information Technology* **24** (1), 4 (2022).
11. D. Newman-Griffis, J. S. Rauchberg, R. Alharbi, L. Hickman, H. Hochheiser, Definition drives design: Disability models and mechanisms of bias in AI technologies. *First Monday* **28** (1), 12903 (2023).
12. M. Langenkamp, D. N. Yue, How Open Source Machine Learning Software Shapes AI. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Association for Computing Machinery, 2022), 385–95.
13. C. Arun, “AI and the global south: Designing for other worlds,” in *The Oxford Handbook of Ethics of AI*, M. D. Dubber, F. Pasquale, S. Das (Oxford University Press, 2019).
14. J. Chubb, D. Reed D, P. Cowling, Expert views about missing AI narratives: is there an AI story crisis? *AI & Society* doi:10.1007/s00146-022-01548-2 (2022).
15. T. M. Porter, *The Rise of Statistical Thinking, 1820–1900* (Princeton University Press, 2020).