

Reflection

For this part of the assignment, you will read Mariette Awad's 2024 / 2025 paper entitled "Types of AI and their use in science."

<https://council.science/wp-content/uploads/2025/10/Types-of-AI-and-their-use-in-science-2.pdf>). Answer the following questions:

1. Awad distinguishes between different "types" of AI. What classification scheme does the paper use and why do these types matter for scientific research?
2. Does Awad make a clear distinction between AI as a tool and AI as a scientific collaborator? If so, What are the differences and what are some examples given to support the differences? Do these examples suggest a real shift in how science is conducted, or mostly an extension of existing methods?
3. What are some limitations or risks of using AI in science? How do these relate to issues such as interpretability, bias, reproducibility, or theory formation?
4. According to Awad's arguments, is AI more likely to accelerate scientific discovery or to reshape the scientific method itself? Do you agree or disagree?

ANSWERS

1. Mariette Awad breaks down AI into three main types depending on their functional applications: foundational paradigms, decision-oriented paradigms, and cross-cutting principles. Foundational paradigms comprise descriptive AI for pattern recognition, predictive AI for predicting trends, and generative AI for generating new content or hypotheses. Decision-oriented paradigms comprise optimization AI for finding optimal solutions and prescriptive AI for prescribing particular actions according to policy objectives. Finally, cross-cutting principles such as privacy-aware, causal, and explainable AI ensure that scientific endeavors are always responsible, transparent, and based on cause-and-effect explanations. These types are important since they enable AI to pursue a wide range of scientific goals, from the early stages of analyzing unlabelled data to the more advanced stages of automating experiments and ensuring scientific integrity.
2. The article makes a clear distinction between AI as a computational tool and AI as an epistemic agent. As a computational tool, AI is used for mundane activities such as data classification or coding assistance using models such as Code Llama. However, as an epistemic agent, AI assists in scientific inquiry by applying statistical and abductive inference to generate explanations based on data patterns rather than theoretical foundations. Examples include AlphaFold 3, which predicts complex biomolecular interactions, and agentic AI models such as LilaSciences or CSIRO's Sciansa, which can plan and prioritize tasks such as a researcher. This is a major paradigm shift in scientific inquiry; computational models are no longer assisting scientific inquiry but are restructuring it by creating knowledge through processes that may defy human intuition.
3. Awad points out that there are major concerns with respect to interpretability and bias. Most of the current state-of-the-art deep learning models are "black boxes," where the output is generated without a clear explanation of the logic behind it, which makes it hard to justify the scientific output. The lack of interpretability is further exacerbated by the issue of bias, where the Natural Language Processing tools can inadvertently encode societal stereotypes about race and gender. With respect to theory formation, models such as Galactica highlight the problem of "hallucinations," where the AI system generates fictional citations or incorrect information that can further undermine confidence in the scientific record. In addition, the issue of reproducibility can arise when the models are highly dependent on a particular dataset or lack

"mechanistic" explanations that enable researchers to understand why errors are being generated.

4. As Awad argues, although AI clearly accelerates discoveries in areas such as drug research and climate modeling, its most significant effect is the transformation of the scientific method itself. This, she argues, marks a "paradigm shift" in the epistemological basis of scientific inquiry. Traditional approaches based on human deductive and inductive reasoning are increasingly supplemented or replaced by statistical and abductive reasoning, in which findings are generated from data patterns rather than from prior theoretical assumptions. As AI systems become increasingly autonomous, they assume functions that were previously the sole province of human beings, such as hypothesis formulation and experimental design. Whether one is in agreement or disagreement with the argument, it is evident from the report that we are headed towards a future in which AI will be a "true generative partner" that demands new ethical frameworks to protect the integrity of human knowledge.