Reflection

Awad distinguishes between different "types" of AI. What classification scheme does the paper use and why do these types matter for scientific research?

In this paper, Awad's classification scheme revolves around different approaches to learning, inference, and knowledge generation. Awad argues that AI should not be understood as a single "monolithic" technology, but a "constellation" of different models, distinguishing them according to their relationship with learning, drawing inferences, and representing knowledge.

Awad identifies the following foundational paradigms: predictive, descriptive, generative, optimization, causal/interpretable, meta-scientific, and privacy-aware AI. Predictive AI forecasts outcomes from data; descriptive AI uncovers patterns and structures; generative AI produces new data or representations; and optimization AI refines experimental conditions to improve results. Causal AI focuses on identifying cause-and-effect relationships, while explainable AI makes model reasoning transparent. Meta-scientific AI assists with hypothesis generation, experimental design, and synthesis. Privacy-aware AI enables analysis of sensitive data without compromising security and privacy.

These distinctions in classification matter in research because each type contributes to scientific research and discovery differently. Some systems, like Predictive AI, increase efficiency and scale, extending and improving established research methods. Others, such as causal/interpretable AI or meta-scientific AI, influence how knowledge is formed, interpreted, and validated. By defining these categories, Awad shows that AI's impact on science depends on the role it plays in the research process. This framework allows the reader to evaluate whether AI is simply accelerating scientific work or reshaping how it is conducted.

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?

Awad makes a meaningful distinction between AI as a research tool and AI as a potential scientific collaborator. She differentiates between systems that enhance human-directed research and those that begin to participate more actively in scientific reasoning and the research process itself.

Many AI systems described in the paper function primarily as tools that extend traditional methods. In this role, AI assists scientists in tasks such as analyzing data, summarizing literature, predicting outcomes, and optimizing experimental parameters. Humans remain responsible for defining research questions, interpreting results, and validating findings. These applications increase speed, scale, and computational capacity without fundamentally altering the structure of scientific inquiry. For example, predictive models used in climate science or genomics accelerate data analysis but do not independently shape research direction. Similarly, optimization models such as deep reinforcement learning in chemical synthesis help identify effective experimental conditions, yet they operate within human-defined goals and constraints.

By contrast, more advanced systems, such as meta-scientific or agentic AI, begin to resemble collaborators. These systems integrate multiple AI capabilities and can support several stages of the scientific process simultaneously, including hypothesis generation, experimental design, analysis, and synthesis. Rather than merely executing predefined tasks, they may propose new hypotheses, suggest research directions, or determine what to explore next. In this sense, AI participates in elements of planning and decision-making. Awad's discussion of automated discovery systems suggests that such models can function similarly to a "junior researcher."

The examples presented in the paper indicate that there is a developing shift rather than a complete one. Even advanced systems operate under human supervision and within human-defined research frameworks, serving largely advisory roles. While AI increasingly contributes to knowledge production by identifying patterns, generating ideas, or automating workflows, it has not replaced human decision-making or responsibility. Therefore, Awad implies that AI not only represents an expansion of existing scientific methods, but the continued development of these models is a beginning to reshape the norms and structure of contemporary scientific practice.

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?

Awad highlights several important limitations and risks that arise as AI systems become more complex and more deeply integrated into scientific research. These risks relate directly to issues of interpretability, bias, reproducibility, and theory formation.

One major limitation is interpretability, as any advanced AI systems, especially deep learning models, act as 'black boxes' whose outputs lack clear explanation. This conflicts with scientific norms that prioritize transparency, and determining causal reasoning. If researchers cannot understand how a model reached its conclusion, it becomes difficult to validate or integrate it into existing theoretical frameworks.

A major risk is that AI systems, particularly NLP and descriptive models, can inherit and amplify biases present in their training data. Awad states that language models may reflect societal biases related to race or gender baked into the underlying datasets. When such models are used in scientific or policy contexts, these biases can distort findings and reinforce existing biases. This raises ethical concerns about fairness and accountability in research.

Reproducibility is another risk, where large language models and other advanced AI systems accelerate research but can be difficult to validate consistently. Like the interpretability concerns, complex "black box" models make it difficult to understand how results are produced, and frequent updates or reliance on proprietary infrastructure further complicate verification. If models are opaque, reproducing results becomes challenging, undermining one of the central principles of science: that findings should be clear, verifiable, and repeatable. Without reproducibility, research cannot be reliably peer-reviewed, errors are harder to detect, and scientific accountability is weakened.

Finally, there are risks to theory formation. AI systems may generate highly accurate predictions or useful hypotheses without explaining underlying mechanisms. If scientific practice becomes overly focused on performance and prediction, there is a risk that explanatory depth

and theoretical understanding will decline in importance. Since AI systems can influence what counts as evidence or which patterns are worth investigating, they may subtly shape the direction of inquiry itself.

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?

Awad suggests that AI is both accelerating scientific discovery and beginning to reshape aspects of the scientific method. She does not present this as a choice or weigh the likelihood of one outcome over the other. Instead, she emphasizes that AI is already speeding research through improved data analysis, modeling, and information sharing, while simultaneously influencing how scientific work is organized and conducted.

At present, AI functions primarily as an accelerator. It increases computational capacity, speeds up experimentation, enhances collaboration, improves efficiency, and enables large-scale data integration. These contributions intensify existing scientific methods rather than replacing them.

However, Awad suggests that AI's impact goes deeper than acceleration. As systems become more autonomous and capable of guiding hypothesis generation, experimental design, and knowledge synthesis, they begin to reshape traditional methodological norms. AI may influence how evidence is evaluated, how research questions are framed, and how scientific information is shared. In this sense, the transformation is gradual but significant.

I agree with Awad's position. Currently, AI acts more as an accelerator, but it has the potential to reshape the scientific process at a deeper level. With the growing complexity of scientific inquiry, researchers are increasingly relying on these tools to manage data, generate insights, and guide experimentation. If AI systems increasingly guide inquiry rather than simply be used as a tool, scientific practice may shift toward a hybrid human/machine model of reasoning. Whether this shift is beneficial will depend on how risks of transparency, bias, and accountability are addressed. As Awad describes the paper itself as a "snapshot," the long-term trajectory of AI in science remains uncertain, especially given the rapid development of AI technologies.