National Network for Critical Technology Assessment

Dashun Wang

Quantifying the Benefits of Artificial Intelligence for Scientific Research

We find AI has been exerting a widespread impact throughout the sciences, but AI education in the current curriculum is not commensurate, and there are demographic disparities in benefiting from AI.

Research Issue

The ongoing artificial intelligence (AI) revolution has the potential to change almost every line of work. As AI capabilities continue to improve in accuracy, robustness, and reach, AI may outperform and even replace human experts across many economically valuable tasks. Despite enormous efforts devoted to understanding AI's impact on labor and the economy, and its recent success in accelerating scientific discovery and progress, we lack a systematic understanding of how advances in AI may benefit scientific research across disciplines and fields. A better understanding of AI's impact on science may not only help guide AI development, bridging AI advances more closely with scientific research but also hold implications for science and innovation policy. This is especially so given AI's recent, remarkable success in advancing research frontiers across several fields, exerting meaningful influence in propelling scientific progress. Building on the growing literature on the future of work and the science of science, here we develop a quantitative framework for estimating AI's impacts on scientific research.

Methods and Data

Our primary dataset contains 87.6 million publications from the Microsoft Academic Graph (1960-2019), spanning 19 disciplines and 292 fields. We integrate this dataset with 7.6 million patents granted by the U.S. Patent and Trademark Office (1960-2019) and the Open Syllabus Project dataset. By applying natural language processing techniques, we measure AI's impact on scientific research at two different levels. Specifically, we first estimate the direct use of AI using an "AI n-gram" framework, where we extract n-grams (bigrams and trigrams) from the titles of AI-related papers and further calculate the weighted frequency of AI n-grams to approximate the direct impact of AI in each field and year. We then estimate the potential impact of AI using a "AI capability field task" framework, where we infer the capabilities of AI by extracting verb-noun pairs from the titles of AI related papers and patents (i.e., what AI can do) and estimate the basic tasks of each field (i.e., what a field conducts) by calculating the relative frequency of verb-noun pairs extracted from the titles of papers published in each field and year. Matching tasks in a field with inferred AI capabilities allows us to approximate AI's potential impact on the field. Moreover, we collected 5.3 million university course syllabi from the Open Syllabus Project (OSP) database and estimated AI education levels in each discipline by calculating the share of references to AI publications. Also, we used the Survey of Doctorate Recipients (SDR) data to solicit scientists' demographic information.

Insights

Our analysis of AI's direct and potential impact on scientific research conveys several findings and sights. First, we find that the use of AI appears widespread throughout the sciences, growing especially rapidly since 2015 (Figure 1a), and papers that use AI exhibit an impact premium, more likely to be highly cited both within and outside their disciplines. Second, despite the heterogeneity of AI's impact across research areas, almost every discipline contains some subfields that benefit substantially from AI innovations (Figure 1b), suggesting AI's pervasive impact across disciplines. Third, analyzing course syllabi across 17 disciplines, we find a systematic misalignment between the education of AI and its impact on research (Figure 1c), suggesting that the supply of AI talents in scientific disciplines is not commensurate with AI research demands. Fourth, the rapid advances pose growing knowledge demands on individual scientists, who increasingly rely on collaborators with AI expertise instead of working solely to push AI applications forward in their disciplines (Figure 1d). Fifth, women and under-represented minorities (URM) scientists benefit substantially less from AI, which may further exacerbate the existing inequalities in science.

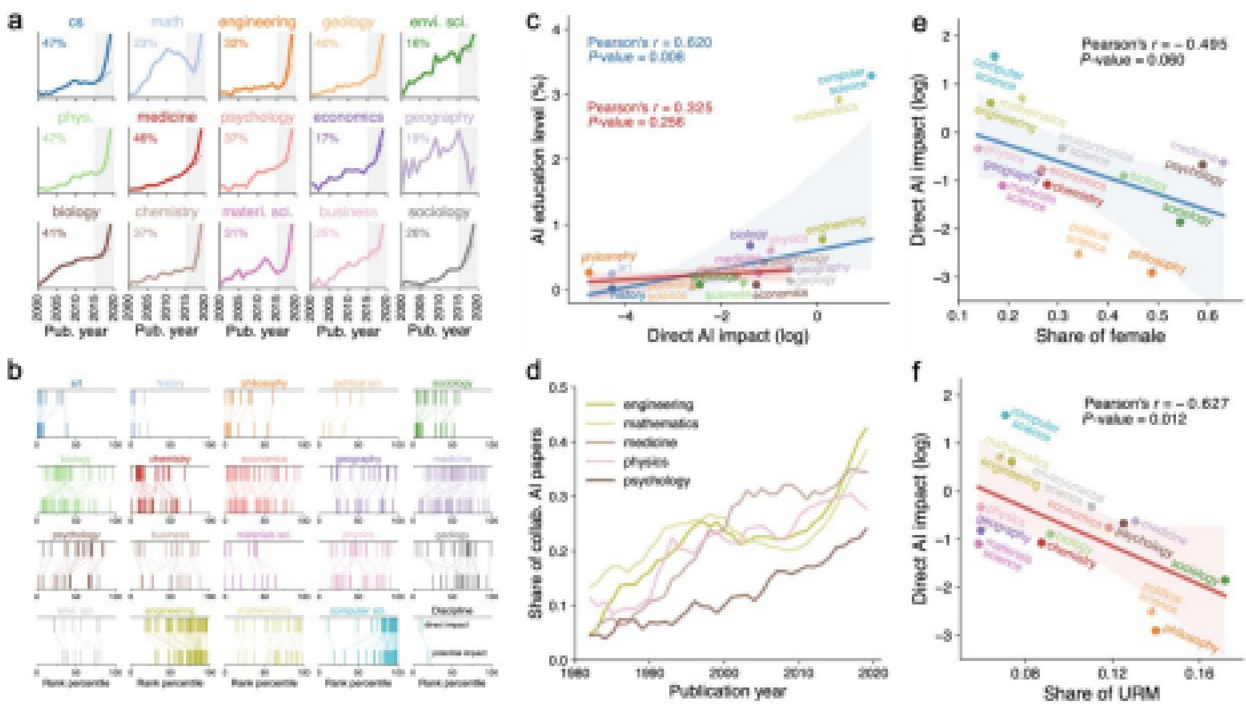

Figure 1. Estimating the benefits of AI in science. (a) Temporal trends of AI's direct impact on different disciplines. (b) The substantial heterogeneity of AI's impact within disciplines. (c) The correlation between the direct AI impact score and the AI education levels. (d) The share of collaborative AI papers across the period of 1980-2019. (e) The negative correlation between AI impact and the share of female scientists. (f) The negative correlation between AI impact score and the share of under-represented minorities scientists.

Options and Trade-offs

The pervasive impact of AI across disciplines and its rapid advances pose growing AI knowledge demands on individual scientists. In particular, the uncovered misalignment between AI's impact on science and the upstreaming AI education indicates a critical need for redesigning current curriculum to teach more AI skills in university or facilitating cross-department collaborations with AI experts. Both education and collaboration on

AI will upskill scientists, which has implications for understanding how to best prepare next-generation scientists who can take full advantage of cutting-edge AI advances for their research. It is also important to recognize that, as AI becomes increasingly capable of performing research tasks, it may create unequal impacts on the research workforce. Our analysis reveals new sources of inequality through AI's benefits for science, which has implications for building and promoting a diverse, equitable, and inclusive research workforce.

Early wins

Our systematic understanding of AI's impact on scientific research as well as the associated heterogeneity may hold implications for informing science and education policy. As AI becomes increasingly capable of performing research tasks, scientists especially those in relevant disciplines will have an opportunity to prepare their AI skills for the future of science. Our findings may also prove useful to the AI community, helping us better understand what capability might be the most fruitful for scientific research. Given that tomorrow's technological developments often begin upstream from scientific research, a better understanding of AI's impacts on science may further inform the range of important policy considerations.