How do AI private sector innovations affect the direction of academic research and career choice?

Mihai Codreanu*, Oscar Aguilar, Jacob Douglas Light, Matheus Dias

Monday 4th September, 2023

Type of project: Research Proposal

1 Challenge addressed / research context

Unlike previous major technological innovations, the rise in AI has been industry-centric. This is due to the easier access of industry to some of the 3 key components of effective LLMs: 1) data, 2) computing power, 3) powerful model architecture. Due to higher capitalization (only new AI firm private investment is about 30x times more than total NSF funding for CS programs[1]), companies tend to completely leapfrog academia in both access to data and access to computing power. As a result, among the 10 most-cited AI papers, only 2 don't have at least one industry affiliated author.[2]

This is important for several reasons. At a basic level, both academia and industry promote innovation but they differ in objective: academia focuses on knowledge production for social welfare, while industry targets knowledge production for private gain. Therefore, the wider social benefit will depend on who produces research, where they produce it, and for what purposes. Additionally, ability to regulate the industry (especially given AI existential risk) may be correlated with performance and understanding of AI technologies of public sector workers. This is all more important since training for AI researchers is publicly subsidized, through generous NSF and public university funding.

*I would like to thank J. Bonney for feedback on this project, and research collaboration on academia-industry ties. Special thanks also for ChatGPT & Bard with help with code & advice this quarter!

[1]Authors' calculations using HAI, Stanford (2023) and NSF (2023).

[2]Authors' ranking and calculations using Google Scholar (2023) citation data.

Using the NSF Earned Doctorates Data, we now document several motivating facts. Over the last decade, computer scientists are increasingly taking jobs in industry. In Figure 1.1, we show the industry growth in levels for current PhD students in Computer Science compared to all PhDs, and Science, non-CS PhDs. The proporion of CS PhDs taking a job in industry increased from less than 50% in 1996, to just over 65% in 2021. In Figure 1.2 we present a different view of the same trend. Here, we see that the growth of CS PhD students in industry is much higher in relative terms than all PhDs, and Science, non-CS PhDs. The growth is measured relative to the 1996 percentages, with CS increasing by almost 20 percentage points, compared to around 15% for all other PhDs and less than 15% for science (excluding CS).

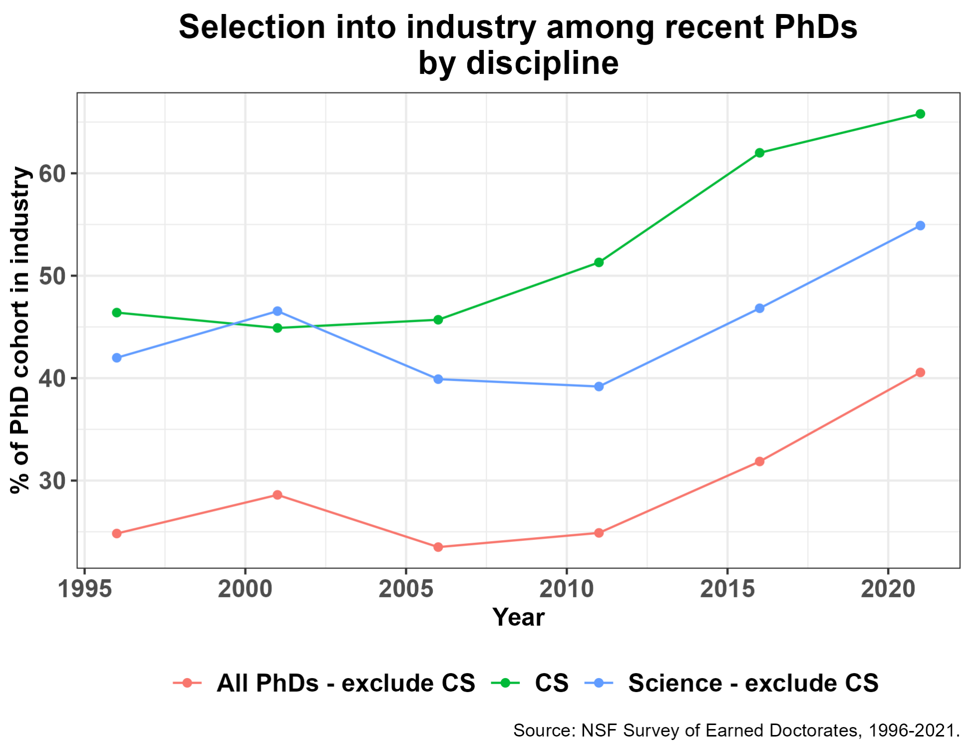

Figure 1.1: Industry Growth Levels 1996-2021

Given this context, one begs the question, what is the role of academia in developing AI technologies? How do students and professors react (in terms of direction of research, techniques used, etc.) to private sector product announcements as well as funding announcements? Do private sector opportunities shocks change (especially top) academic career trajectories? What does that mean for society?

In this research proposal, we will be able to offer a framework for a first go to answering these questions, and outline useful variations in the timing of innovations and the general public announcements, incentives to innovate, and local private sector opportunities. To do

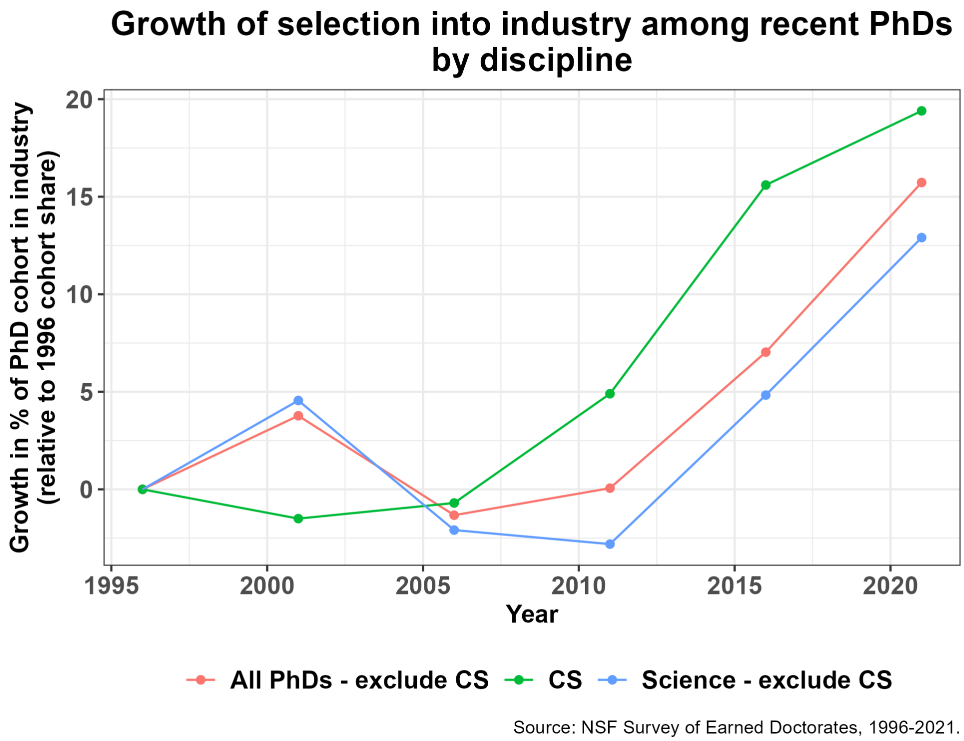

Figure 1.2: Industry Growth Relative to 1996

so, we will plan to use extensive GitHub anonymized data from companies and academic institutions (obtained from a private relationship with Microsoft), a dataset of publication records across different disciplines, scraped information on graduate students and professors using academic website, student placement records and declared private sector affiliations for professors. We will use: event study analyses, difference-in-differences approaches and more complex IV and potentially structural models for uncovering these effects.

2 Team

We build this team because: i) we were all interested in writing a research proposal rather than focusing on a business/policy plan, ii) we all have interests in higher education and the relationship between firms and academia, as well as the effects of AI on students' and professors' research and universities, iii) we have different interests and background (Econ of innovation–Mihai, Econ of education–Jacob, Statistics–Matheus, Management Science Engineering–Oscar), as well as various exposure to the private sector (working at start-ups, unicorns, big tech). Mihai coordinated the team, and put together the idea, proposal, and presentation, Jacob helped with motivating the project, structuring and streamlining the research area and some data work, Oscar (with impressice work on the arXiv dataset) and Matheus (scraping) did the heavylifting in terms of data, and used their private sector experience in streamlining the research and questions to look at. We all collaborated with our diverse skills on making this a success and enjoyed the process!

3 Research Questions & Literature Review

This paper primarily tackles the research question: how does AI innovation within the private sector influence the trajectory of academic research and professor and grad students career choices? In order words, does the progress in private sector innovation alter or "crowd out" the research conducted in academic settings? To answer these question, we will also dig into academia's role in the evolution of AI technologies and conclude investigating the social welfare implications of an increasing emphasis on privately conducted AI research.

This study builds on several strands of literature.

The first branch explores the choice to become a scientist and the drivers of shifting career priorities. Bell et al. (2019) illustrated the impact of labor market conditions and public tax policies on long-term career outcomes for scientists, while Bianchi and Giorcelli (2020) analysed the long-term effects of STEM education exposure during graduate training. Biasi and Ma (2023) further discuss the influence of academia's course offerings on the career paths of new Ph.D. graduates in US universities. Myers (2020) investigates how early-career experiences shape lifetime research trajectories for scientists and how changes in funding impact the elasticity of researched topics. Finally, Sauermann and Roach (2012) explore the motivations leading scientists to enter and persist in careers with either academic or industry emphasis. Our study extends these findings by examining the causal impact of specific private sector innovations on researchers' decisions in the context of AI.

The second literature strand looks at the role of competition in the effectiveness of scientific research. Azoulay et al. (2019) examine the impact of the passing of star scientists on the subsequent creativity of researchers working in the same field. Wang et al. (2017) study how publishing pressure affects the novelty of scientific production and propensity for article retraction. Our contribution to this literature is to focus on a different dimension of competition, specifically competition from private sector researchers and the potential inequality of (monetary) resources.

Finally, the third literature area analyzes societal welfare implications and long-term equitable growth from technological progress. Romer (1990) first introduces endogenous technological change into growth models. Aghion and Howitt (1992) further develop Schumpeterian growth theory to explain the observed persistence of growth rates. Acemoglu (2002) formalizes the theory of endogenous technological change, and Jones (2021) explores the economic mechanisms that drive wealth concentration and its implications for long-term economic growth. This is also related to David (1985) which investigates the role of technological change and decisions on long-term outcomes and norms. We augment this literature by analyzing the welfare effects of a new wave of technological innovation, namely the AI "awakening".

4 Data and Methods

Our study draws data from a range of short-term and long-term sources. For short-term data, we will primarily rely on the following:

1. **GitHub data on a subsample of companies and individual anonymized research code**. This is a small random sample from the population of GitHub users, which has high penetration, hopefully provided by our contact at Microsoft GitHub. This includes data on the type of packages and algorithms they use, their code writing timelines, measurements of AI written code, and intra-institution or intra-company collaborations. This data was not obtained yet, but we are in contact with a GitHub Research employee to make this happen soon.

2. **Additional GitHub data scraped from personal academic websites, complemented by data from CVs, and Linkedin**. This extra data set is meant to supplement the data provided by Microsoft, specifically for academics who share more identifiable information. This also serves as a source of controls and heterogeneity analysis.

3. **Announcements by companies about funding, product releases, and scientific developments**. This data is scraped from company websites, social media platforms like Twitter, and aggregator websites such as Jack Clark's Import AI.

4. **Published working papers from arXiv**, publications from journals and databases **(Scopus, Web of Science, JSTOR)**.

For long-term data, we will use the following:

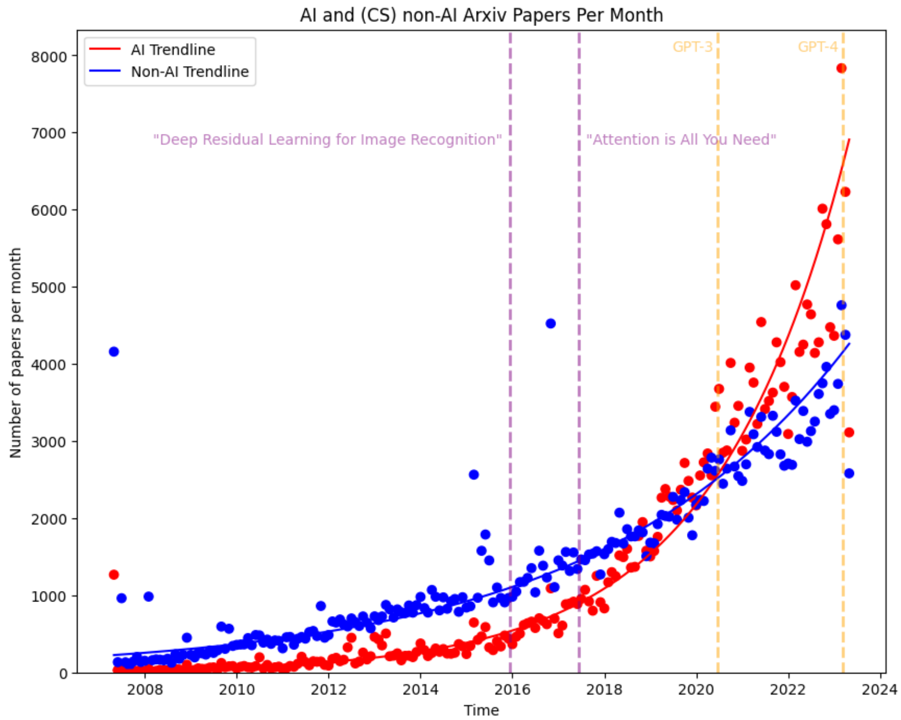

Figure 4.1: Exponential growth in AI vs. non-AI research publications on arXiv

1. **University-level data on student placements and on collaborations between academia and industry**. This data will be obtained from official university sources or scraped. In this proposal, we will use a subsample from Stanford.

2. **Complementary data on a student future research** (if they remain in academia), which is sourced from academic websites. If the student ventures into entrepreneurship, we will **collect data on any new businesses started from the respective company websites**. In the case of a position in an established firm, we source information from LinkedIn profiles and company websites using disambiguation algorithms.

As a first go, we analyze data from arXiv, an open-access repository of scientific papers, which is very popular in Computer Science. We want to see to what extent developments in AI are captured by this source of data. Figure 4.1 reveals an exponential growth in the number of papers published in the domain of AI and related topics. The growth rate of AI research significantly surpasses that of non-AI computer science-related research, with an inflection point around the publication of the "Attention is All You Need" ('Transformers' paper).

The question arises whether this growth is driven by new entrants into the field. We then focus on a subsample that we scrape, of current academics and students affiliated with

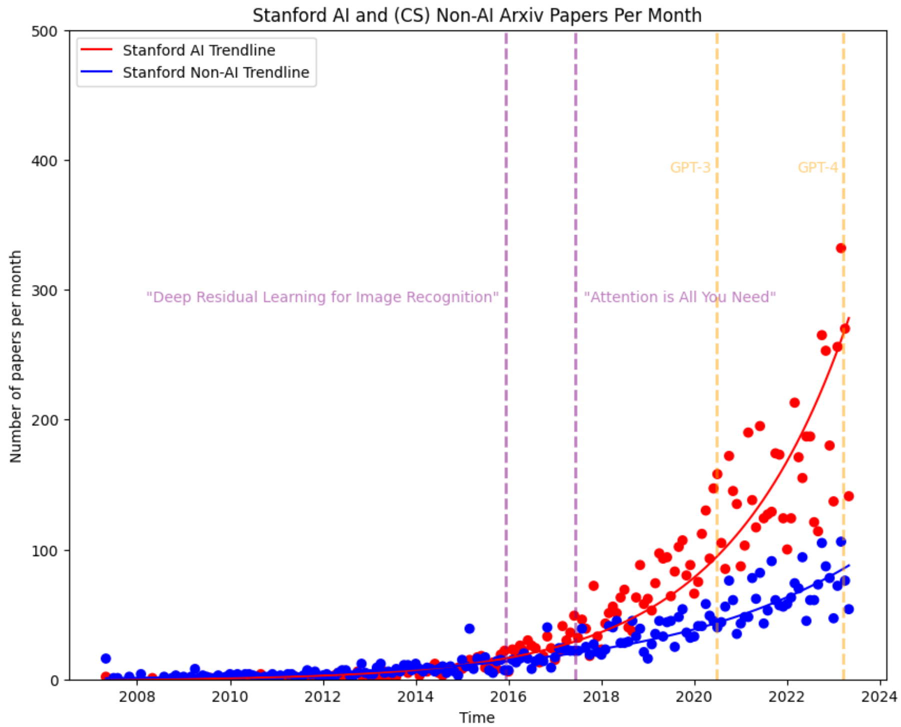

Figure 4.2: Growth in AI vs. non-AI research publications from current Stanford affiliates

Stanford University, classifying their work since 2007 irrespective of the affiliation at the time of publication. Figure 4.2 shows that there is also exponential growth in Stanford's AI research; it is interesting to note that the acceleration in Stanford AI v. non-AI research seems to predate that of the overall sample.

Is there any evidence that the focus on new AI methodologies crowding out older methods? In Figure 4.3 we examine this by focusing on the Natural Language Processing (NLP) subdomain and comparing the research trends between "emerging" methods (abstracts that contain keywords such as contextual embeddings[3]) and "declining" methods (such as bag of words[4]). As of now, we observe no changes in the trends for research on conventional, "declining" methods.

So far, we have established various data sources, and some interesting facts using preliminary arXiv and Stanford scraped data. We now proceed to outline the next steps in terms of the methods that we will use. Our methodology for examining the impact of advancements

[3] The full list is: transformers, contextual embeddings, capsule networks, zero-shot learning, few-shot learning, transfer learning, neural machine translation, dialogue systems, chatbots, cross-lingual models and large language models.

[4] The full list is: bag of words, TF-IDF, rule based systems, feature engineering, N-grams, latent semantic analysis, part of speech tagging.

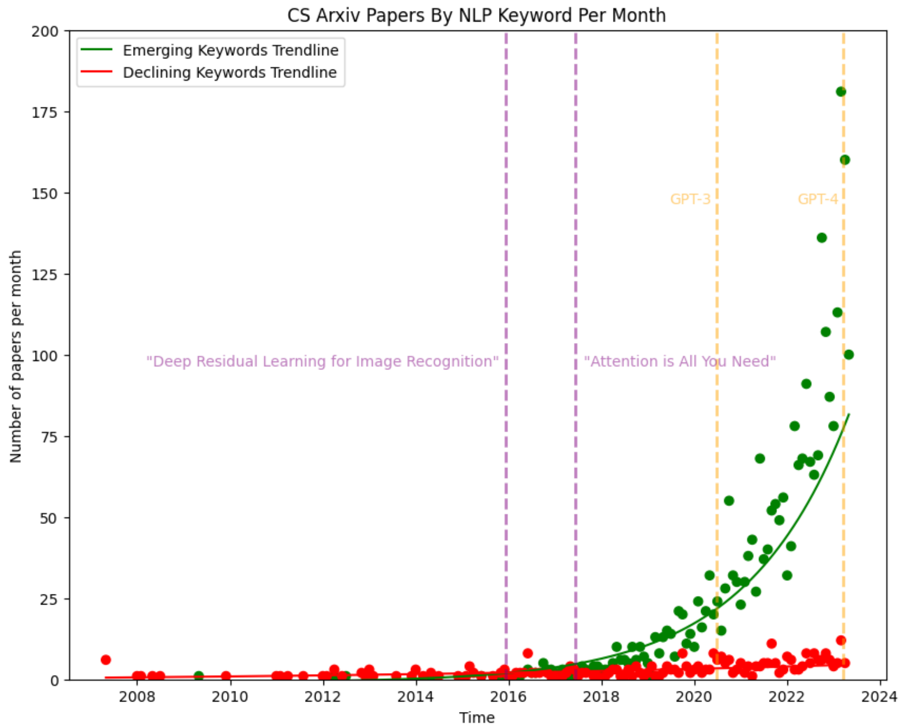

Figure 4.3: Trends in research on "emerging" vs. "declining" NLP methods

in AI on research trends leverages three analytical frameworks, each addressing different aspects of the research objectives.

Firstly, we aim to use event study analyses to identify the short-term effects of pivotal moments in the AI technology landscape, such as the release of the Transformer paper, the launch of ChatGPT, and Google Bard announcements. The objective is to observe the short-term reactions of graduate students, professors, and private firms. Specifically, we will focus on how these groups adjust their ongoing projects and the coding they employ in response to these advancements. The time frames for these studies are expected to span a few months around each key event. The first estimate can be represented by the equation:

$$Y_{it} = \alpha + \sum \beta_t \times Pre/Post_t + \gamma_i + \lambda_t + \varepsilon_{it} \tag{1}$$

where Y_{it} is the dependent variable for entity i at time t. It can represent various dimensions like coding work volume, research direction (quantified as projects worked on in an area). $Pre/Post_t$ is a set of dummy variables, each of which equals 1 for a specific period, daily or monthly before and after the event, and 0 otherwise. γ_i denotes entity fixed effects for entities like a firm, institution, or author, controlling for their time-invariant characteristics. λ_t stands for time fixed effects, which control for overall trends that affect all entities, such as a general increase in AI publications. The coefficients β_t estimate the average effect of the event on the outcome variable at each period after, relative to the period just before.

Secondly, we propose to undertake a Difference-in-Differences (DiD) analysis, identifying groups differentially affected by these seminal papers or developments in the AI field. The process will necessitate a definition for innovations (which we may take from e.g. ImportAI or define based on e.g. improved model performance). The comparisons we plan to make could involve Computer Science (CS) researchers in AI/ML vs. other CS disciplines, CS PhDs vs. other scientific fields, or professors in the period just before and after receiving tenure (which changes the incentives for academic work, keeping commercialization/industry incentives roughly constant). The time frames for these analyses could range from a few months to a few years, depending on the specific comparison. We use the equation:

$$Y_{ijt} = \beta_0 + \beta_1 \times Post_t \times Treat_i + \sum \delta_t \times Year_t + \gamma_i + \lambda_j + \varepsilon_{ijt} \tag{2}$$

Here, Y_{ijt} is the dependent variable for entity i in sector j at time t. It, too, can represent dimensions like coding work volume, research direction, or career choice. $Post_t$ is a dummy variable that equals 1 for all periods after the private sector announcement and 0 otherwise. $Treat_i$ is a dummy variable that equals 1 for the treated group, which may consist of AI students potentially influenced by the announcement, and 0 for the control group which could be composed of CS or Engineering students likely not affected by the announcement. $Year_t$, γ_i, and λ_j are fixed effects controlling for time, entity, and sector (/industry) time invariant characteristics. The key assumption for this model is that there are parallel trends and no anticipation effects.

Thirdly, we plan to employ Instrumental Variable (IV) and more advanced structural approaches to address questions concerning work placement and welfare, which cannot be adequately answered with less complex techniques. While we are still considering the implementation details of these methods, we are contemplating using local labor market conditions related to AI opportunities as an instrument. This approach presents challenges due to factors such as the increasing prevalence of remote work.

5 Limitations, Future Research and Conclusion

Our approach has several limitations, some of which are inherent to the complexity and cutting-edge nature of AI research. First, it is challenging to measure research directions beyond general areas (e.g. AI, ML, NLP, etc.). In the GitHub data, we will be distinguishing between e.g. packages and coding languages used, but going beyond will require a detailed understanding of computer science, as well as the ability to keep up with rapidly evolving AI and ML techniques. Additionally, we will need to access full text (and maybe code) of the publications, not just abstracts, which may present challenges related to data processing.

The event studies methodology, while powerful in isolating immediate effects of key events, focuses on short-term responses and might miss important longer-term shifts in behavior and research direction. Our initial findings suggest that there aren't dramatic immediate effects observable in the volume of AI publications on arXiv following key AI advancements, and the effects are in second order trends changes i.e. more subtle.

The Difference-in-Differences approach presents its own unique set of challenges. The interpretation of coefficients may be complicated due to the pervasive impact of AI across various fields and disciplines. Finding 'good' groups that are completely unaffected by AI developments for comparison might be difficult. The potential ubiquitous nature of AI's influence could lead to bias in our estimates if our 'control' groups are also indirectly affected by the AI advancements we study.

The use of IV and more complex structural methods, while potentially offering deeper insights into the questions of work placement and welfare, could be challenging to implement. In particular, we envision using local labor market conditions (e.g. job availability/postings) related to AI opportunities as instruments. However, the increasing prevalence of remote work complicates the interpretation and measurement of 'local' labor market conditions. This aspect might necessitate the development of new measures to capture the digital or remote 'locality' in the AI job market, which could introduce additional complexities and measurement error into our analysis.

However, notwithstanding these difficulties, we believe that this research not only offers an engaging examination from a causal inference perspective, but also a valuable descriptive study. It enables us to provide more granular insights into the work and research direction of academics in the age of AI. Furthermore, this research serves as an important case study that aids our understanding of the broader implications of industry-academia collaborations, potentially informing policy, funding decisions, and academic initiatives in the rapidly transforming AI and non-AI landscape. This kind of detailed exploration of the academic world in the era of AI can shed light on the broader changes in our society and economy, triggered by these highly transformative technologies. We are excited to keep working on this project!

References

Acemoglu, D. (2002). Directed technical change. *The Review of Economic Studies*, 69:781–809.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60:323–351.

Azoulay, P., Fons-Rosen, C., and Zivin, J. S. G. (2019). Does science advance one funeral at a time? *American Economic Review*, 109:2889–2920.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Reenen, J. V. (2019). Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics*, 134:647–713.

Bianchi, N. and Giorcelli, M. (2020). Scientific education and innovation: From technical diplomas to university stem degrees [the social origins and iq of inventors]. *Journal of the European Economic Association*, 18(5):2608–2646.

Biasi, B. and Ma, S. (2023). The education-innovation gap.

David, P. A. (1985). Clio and the economics of qwerty. *The American Economic Review*, 75(2):332–337.

HAI, Stanford (2023). Measuring trends in artificial intelligence.

Jones, C. I. (2021). Taxing top incomes in a world of ideas. *The Review of Economic Studies*, 88:1545–1575.

Myers, K. (2020). The elasticity of science. *American Economic Journal: Applied Economics*, 12:103–134.

NSF (2023). NSF Funding Figures. National Science Foundation (NSF).

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98:71–102.

Sauermann, H. and Roach, M. (2012). Science PhD career preferences: Levels, changes, and advisor encouragement. *PLOS ONE*, 7(5):e36307.

Wang, J., Veugelers, R., and Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436.