Is there a narrowing of AI research?

J. Mateos-Garcia, National Endowment for Science, Technology and the Arts, United Kingdom

J. Klinger, National Endowment for Science, Technology and the Arts, United Kingdom

Introduction

Large technology companies have largely driven recent advances in artificial intelligence (AI) by developing or deploying deep-learning techniques at scale. This essay examines the current state of play in AI research, including rationales for preserving technological diversity. It subsequently identifies two mutually reinforcing economic processes that may work to reduce this diversity: economies of scale and scope, and collective choice problems. The role of the private sector in further narrowing diversity is also explored. The essay ends with suggestions on how policy makers could promote technology diversity in AI research from both supply and demand sides.

Recent advances in AI have in great part been driven by deep-learning techniques developed and/or deployed at scale by large technology companies. DeepMind, a research lab owned by Alphabet, has produced important breakthroughs in game-playing and protein structure prediction. Google has developed key techniques for language modelling such as word2vec and BERT. Microsoft built the first speech recognition system to reach human-level performance. OpenAI, a not-for-profit institute backed by Microsoft, has created and commercialised GPT-3, a large language model with powerful text generation capabilities (see the essay in this report by Jungwon Byun and Andreas Stuhlmüller). The most popular software frameworks for AI research and development (R&D) – TensorFlow and Pytorch – are maintained by Google and Facebook, respectively.

Many of the ideas underpinning these advances originated in academia and public research labs, suggesting an effective flow of knowledge from the public sector to industry. Commercial demand for AI graduates is also at an all-time high (Jurowetzki et al., 2021). Meanwhile, researchers in universities and the public sector are increasingly adopting powerful open-source software tools and models developed in industry.

However, the short-term benefits of rapid advances in deep learning and the tighter intertwining of public and private research agendas is not without risks. Indeed, a growing number of scientists and technologists has expressed concerns about the possible downsides of data and compute-intensive deep-learning methods that have come to dominate AI research. They also point to an excessive influence of corporate interests in the trajectory of such research (Marcus, 2018; Bender et al., 2021; Whittaker, 2021). Thus, the question arises: is AI research becoming too focused on a narrow set of ideas and methods aligned with the interests of influential corporate players?

In “A narrowing of AI research?”, Klinger et al. (2020) address this question by measuring the thematic diversity of the topics studied by AI researchers. They look at how topics have evolved over time, compare the diversity of AI research in academia and industry, and explore the influence of private sector research (proxied via citations) in the evolution of the field. Results suggest that technological diversity in AI research has stagnated in recent years. In addition, leading, highly influential private sector companies tend to focus on a narrower set of state-of-the-art methods and techniques than universities. This could provide a rationale for policy interventions to preserve diversity in AI research.

The current state of play

Rationales for preserving technological diversity in AI research

It is often possible to achieve the same practical goal through different technological designs. For instance, an automobile can be powered by a combustion engine or an electric motor. Similarly, an AI system can base its decisions on a collection of logical rules or on a machine-learning model. A plurality of methods can be deployed to produce scientific knowledge in the same domain.

There are several reasons why it may be desirable to preserve such technological diversity: creativity, inclusiveness and resilience.

Creativity

Innovation involves the creative recombination of ideas, and unusual mixes are often an important source of radical and transformative innovations (Arthur, 2009). For example, today's deep-learning methods emerged at the intersection of computer science and neuroscience. Some recent advances in AI such as the AlphaGo program that defeated Go world champion Lee Sedol in 2016 also brought together state-of-the-art deep reinforcement learning techniques and traditional tree-search algorithms (Pumperla and Ferguson, 2019). Many researchers believe it is possible to overcome some limitations in deep learning-based AI systems by combining techniques from other AI traditions such as symbolic logic, causal inference or intelligence augmentation (Pearl, 2018; Marcus and Davis, 2019). A more homogenised, less diversified landscape of AI research will contain a less varied set of ideas that could be recombined in this way, potentially decreasing innovation and hindering attempts to overcome the limitations of today's dominant deep-learning designs.

Inclusiveness

Only rarely will a single technology be equally suitable for all applications, sectors and communities. In the case of AI, some sectors such as advertising and media are awash with user data that can be used to train and target deep-learning models. However, other sectors, such as education, are less data-intensive.[1] The high-stakes nature of decisions in the health sector renders deep learning less suitable than in social media or search applications. This is due to the "black-box" nature of deep-learning systems, which makes the algorithmic processes driving their results less open to direct inspection (Miotto et al., 2018; Marcus and Davis, 2019). This means that a loss of technological diversity in AI could lead to some sectors or communities lacking AI systems adapted to their needs and contexts.[2]

Resilience

Homogeneous technological ecosystems (monocultures) are more vulnerable to changes in circumstances, including the discovery of unexpected defects or limitations in a dominant design. In those cases, problems could arise from not having preserved alternative technologies that could readily be adopted. For example, the depletion of global oil reserves and the recognition of the environmental impact of CO_2 emissions called into question the reliance on combustion engines. In the case of deep learning, there is increasing evidence that AI systems based on these techniques have various weaknesses. They tend to be brittle with limited ability to generalise outside of the datasets they were initially trained on. They are also vulnerable to gaming by malicious users. In addition, they could have substantial environmental impacts due to their reliance on energy-intensive computation. Systems based on deep learning can also be unfair. Since they have to be trained on large datasets, it is sometimes uneconomical to carefully filter out biased, prejudiced and/or inflammatory input data that could skew their outputs (Strubell, Ganesh and McCallum, 2019; D'Amour et al., 2020; Raji et al., 2021).

Reasons to expect a narrowing in AI research

There are two mutually reinforcing economic processes that may work to reduce technological diversity in AI research: economies of scale and scope, and collective choice problems.

Economies of scale and scope

In these situations, an increase in the supply or adoption of a technology makes it more attractive than its competitors. One important example of scale economies is network effects, where the size of a technology's user base increases its value to subsequent users. Network effects can potentially lead to situations where random fluctuations in adoption levels of a technology might tip the market in its favour independently from the technology's objective quality (Arthur, 1994). In two-sided markets, a technology's attractiveness depends on the presence of complementary assets such as data, computational infrastructure and/or skills. Such markets are especially important for ICT systems such as AI (Rochet and Tirole, 2006). This is demonstrated by several episodes in the history of AI where a technique benefited from independent improvements in complementary technologies that made it easier to adopt (Hooker, 2020). For example, the arrival of graphics processing units (GPUs) that could render computer game graphics efficiently lowered the barriers to deploying compute-intensive deep-learning techniques. Without the arrival of GPUs, other AI techniques may have prevailed. Once a technological design gains an edge over its competitors, this creates incentives to invest in complementary resources that strengthen that technology's advantages.

Collective choice problems

The uncoordinated behaviours of individual actors such as research teams or firms could limit technological diversity. For example, innovators could have fewer incentives to invest in developing a second-tier technology against a dominant one (Acemoglu, 2012). They may also choose to focus on technologies that generate more short-term returns, even when they know alternatives would be more beneficial in the longer term (Bryan and Lemus, 2017). In the case of AI, there is a growing sense that publication, commercial and geopolitical races could be encouraging such short-term behaviours (Armstrong, Bostrom and Shulman, 2016). Research teams and countries competing fiercely to advance the state-of-the-art in AI benchmarks, launch new products and become global AI leaders are more likely to focus their efforts on advancing the dominant paradigm (i.e. deep learning). They are less likely to explore "second-tier" techniques that preserve AI's technological diversity but have uncertain benefits.

The role of the private sector

Private sector participation in the development of a technology could intensify the pressures that narrow it. After all, commercial actors have strong incentives to invest in technologies that can be more readily deployed and to leverage their investments in technology across more markets. The results can include behaviours that drive product life cycles in certain ways. The exploration of alternative technologies, for example, could be followed by exploitation of a dominant design (Utterback and Abernathy, 1975); a competition to establish technical standards that harness network effects to dominate the market (Shapiro and Varian, 1998); and homogenisation of industries as organisations become more similar to each other to facilitate flows of knowledge and talent (Beckert, 2010). Businesses might also steer the trajectory of a technology in directions aligned with their particular interests, potentially neglecting negative externalities, unintended consequences and societal preferences.

These concerns are visible in AI research as large technology companies become more influential. These firms are making vast investments to develop deep-learning techniques that complement their assets (big data and computational infrastructure) and applications (e.g. information search, content filtering and ad-targeting). Some evidence suggests these investments are draining researchers from academia. Similarly, evidence points to skewed research priorities of public research labs that receive private funding from and/or need to collaborate with industry to access the large datasets and infrastructures required for cutting-edge research (Jurowetzki et al., 2021; Whittaker, 2021). Meanwhile, technology companies might have incentives to downplay the limitations and risks of deep-learning techniques that increasingly sit at the core of their products and services (Bender et al., 2021). All of this could lead to what some researchers have referred to as a "de-democratisation" of AI research. In such an environment, AI research focuses on a narrow set of compute-intensive techniques mainly developed and deployed by a small number of private research labs and their collaborators in elite universities (Ahmed and Wahed, 2020).

A further narrowing of AI research?

The discussion above has provided theoretical reasons and supporting evidence for three points:

1. It would be desirable to preserve AI's technological diversity.
2. Economies of scale and scope and collective choice problems could make AI research narrower.
3. Increasing private sector participation and influence in AI research may intensify this process.

The paper, "A narrowing of AI research?" (Klinger et al. (2020), sought to improve the evidence base about points (2) and (3). It conducted a quantitative analysis of 1.8 million articles from *arXiv*, a preprint repository widely used by the AI research community to disseminate its work. Having identified around 100 000 AI papers in this corpus, the authors analysed their abstracts to measure thematic concentration and heterogeneity and construct several indicators of technological diversity.[3] They then analysed the evolution of these metrics over time, thus addressing point (2).

They also extracted information about the institutional affiliation of each article's authors. This aimed to measure private sector participation in AI research. It also aimed to compare the thematic diversity and influence of "public" and "private" sector AI research overall, thus addressing point (3). Three key findings are summarised below.

There is evidence of a recent stagnation and even decline in the diversity of AI research

All metrics of diversity show that technological diversity in AI research has expanded since the late 2000s. However, this growth has stagnated and even started to decline from the mid-2010s. This is despite a substantial increase in the number of AI publications in recent years (60% of the AI articles in the corpus studied were published after 2018). Such an increase might have been expected to expand the range of AI techniques and applications explored. Analysis of the factors behind the stagnation of technological diversity in AI research shows increasing concentration of research in a small number of influential topics related to deep learning.

Private AI research is thematically narrower and more influential than academic research, and it focuses on computationally intensive deep-learning techniques

Private companies are ten times more likely to participate in AI research than in other research contained in *arXiv*. In 2020, 20% of AI papers involved at least one researcher affiliated with a private company, with large US technology companies such as Google, Microsoft, IBM, Facebook and Amazon ranking highest.

This private body of AI research is narrower than the public body according to all the metrics used, even after adjusting for differences in the number of papers produced. The analysis also shows that private companies tend to have narrower research profiles than universities after controlling for their volume of AI research, the year a paper was published and unobservable organisation-specific factors. Private companies tend to specialise in state-of-the-art deep-learning topics in computer vision and computer language; infrastructure to scale up computationally intensive AI methods; and applications in online search, social media and ad-targeting. By contrast, they tend to be less focused on health applications of AI and analyses of the societal implications of AI.

The authors also found that AI research involving companies tends to be more highly cited even after controlling for its topic. This is consistent with the idea that the private sector might be shaping the evolution of the field directly through the research it publishes, and indirectly by providing a foundation that other researchers build on.

Elite academic institutions have similar research profiles to private sector institutions

Some of the largest and most prestigious universities have lower levels of thematic diversity in AI research than would be expected given their volume of activity and public nature. These institutions include MIT; University of California, Berkeley; Carnegie Mellon; and Stanford University. Such influential universities tend to be the top collaborators of private companies, suggesting some homogenisation at the top of AI research.

Conclusion

The analysis has some limitations given the authors only consider published AI research. They were not able to make any causal statements about the direct impact of private sector participation in AI research. Nor could they make strong inferences about why companies are less thematically diverse than other institutions.

Perhaps most importantly, the analysis says little about the impacts of a loss of technological diversity in AI research. The concentration of efforts in a dominant design evidenced here might simply be making research more efficient by reducing the dissipation of efforts down unproductive dead-ends. Some theoretical and qualitative arguments were made for why this perspective might be excessively optimistic, but more evidence is required to bolster this case. This would require quantification of the loss of resilience, creativity and inclusiveness brought about by a narrowing of AI research – all of which are important questions for future work. Lacking that, portfolio theory suggests likely disadvantages in focusing all (or most) AI research on a single family of (deep-learning) techniques that, as a growing number of voices in the field argue, also have important limitations and risks. This provides a rationale for policy makers to consider how they might spur technological diversity in AI research.

What can policy makers do?

Technological diversity and the supply side

The analysis shows that universities tend to produce more diverse AI research than the private sector. Accordingly, bolstering public R&D capabilities might make the field more diverse. This could be done through increases in the levels of research funding and the supply of talent, computational infrastructure and data for publicly oriented AI research. A larger talent pool would reduce the impact of a brain drain of AI researchers from universities to industry. Better public cloud and data infrastructures would also make academic researchers less reliant on collaboration with private companies (Ho et al., 2021).

Recognising the propensity towards "research bandwagons" in academia, research and funders should pay special attention to projects that explore new techniques and methods separate from the dominant deep-learning paradigm. This may require patience and a tolerance of failure. Initiatives to increase the sociodemographic diversity and inclusiveness of the AI research talent pool would broaden the range of perspectives and preferences brought into the design and evaluation of AI technologies. This might make them more thematically diverse (Acemoglu, 2012).

AI researchers in the public sector have much to learn from industry teams that regularly share their code and data and build robust tools to make their findings easier to reproduce and their methods easier to deploy. Policy makers should strengthen incentives for adoption of these open science and open-source methods in academia.

Demand side of technological diversity

New datasets, benchmarks and metrics could reflect the limitations of deep-learning techniques (for example in terms of energy consumption) and the advantages of their alternatives. In so doing, they could help steer the efforts of AI research teams. Mission-driven innovation policies could encourage deployment of AI techniques to tackle big societal challenges and increase adoption of AI methods in underserved sectors. This, in turn, could spur development of new techniques more relevant for domains where deep learning is less suitable. Finally, policy makers could consider regulatory interventions, possibly focused on specific use cases, that penalise the negative externalities of deep-learning methods. One case, for example, is the impact of environmental costs and risks for minority communities. Such interventions might encourage their developers and adopters to explore alternative techniques.

Designing and implementing these initiatives will require policy makers to overcome three substantial barriers:

1. There are strong incentives for policy makers to retain a short-sighted focus on exploiting dominant AI technologies. Instead, they need to explore alternatives relevant to their countries, societal challenges, and scientific and technological capabilities.
2. Policy makers needs more expertise and know-how to help them decide what sort of technology initiatives to support.
3. Policy makers need to countervail the massive investments by the private sector on AI R&D. Their ability to do so could help prevent a premature lock-in to powerful yet limited deep-learning techniques. This could provide the foundations for future AI revolutions with fewer risks and more widely shared benefits.

References

Acemoglu, D. (2012), *Diversity and Technological Progress,* University of Chicago Press.

Ahmed, N. and M. Wahed (2020), “The de-democratization of AI: Deep learning and the compute divide in artificial intelligence research”, *arXiv,* preprint arXiv:2010.15581, https://doi.org/10.48550/arXiv.2010.15581.

Armstrong, S., N. Bostrom and C. Shulman (2016), “Racing to the precipice: A model of artificial intelligence development”, *AI & Society*, Vol. 31/2, pp. 201-206.

Arthur, W.B. (2009), *The Nature of Technology: What It Is and How It Evolves,* Simon & Schuster, New York.

Arthur, W.B. (1994), *Increasing Returns and Path Dependence in the Economy,* University of Michigan Press.

Beckert, J. (2010), “Institutional isomorphism revisited: Convergence and divergence in institutional change”, *Sociological Theory*, Vol. 28/2, pp. 150-166, www.jstor.org/stable/25746221.

Bender, E.M, et al. (2021), “On the dangers of stochastic parrots: Can language models be too big?”, in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, https://doi.org/10.1145/3442188.3445922.

Bryan, K.A. and J. Lemus (2017), “The direction of innovation”, *Journal of Economic Theory*, Vol. 172, pp. 24772, https://doi.org/10.1016/j.jet.2017.09.005.

D'Amour, A. et al. (2020), "Underspecification presents challenges for credibility in modern machine learning", *arXiv*, preprint arXiv:2011.03395, https://doi.org/10.48550/arXiv.2011.03395.

Ho, D.E. et al. (2021), *Building a National AI Research Resource: A Blueprint for the National Research Cloud,* The Stanford Institute for Human-Centered Artificial Intelligence, Stanford, https://hai.stanford.edu/sites/default/files/2021-10/HAI_NRCR_2021_0.pdf.

Hooker, S. (2020), "The hardware lottery," *arXiv*, preprint arXiv:2009.06489, https://doi.org/10.48550/arXiv.2009.06489.

Jurowetzki, R. et al. (2021), "The privatization of AI research (-ers): Causes and potential consequences – from university-industry interaction to public research brain-drain?" *arXiv*, preprint arXiv:2102.01648, https://doi.org/10.48550/arXiv.2102.01648.

Klinger, J. et al. (2020), "A narrowing of AI research?", *arXiv*, preprint arXiv:2009.10385, https://doi.org/10.48550/arXiv.2009.10385.

Marcus, G. (2018), "Deep learning: A critical appraisal", *arXiv*, preprint arXiv:1801.00631, https://doi.org/10.48550/arXiv.1801.00631.

Marcus, G. and E. Davis (2019), *Rebooting AI: Building Artificial Intelligence We Can Trust,* Vintage, New York.

Miotto, R. et al. (2018), "Deep learning for healthcare: Review, opportunities and challenges", *Briefings in Bioinformatics*, Vol. 19/6, pp. 1236-1246, https://doi.org/10.1093/bib/bbx044.

Pearl, J. (2018), "Theoretical impediments to machine learning with seven sparks from the causal revolution", *arXiv*, preprint arXiv:1801.04016, https://doi.org/10.48550/arXiv.1801.04016.

Raji, I.D et al. (2021), "AI and the everything in the whole wide world benchmark", *arXiv*, preprint arXiv:2111.15366, https://doi.org/10.48550/arXiv.2111.15366.

Rochet, J.-C. and J. Tirole (2006), "Two-sided markets: A progress report", *The RAND Journal of Economics*, Vol. 37/3, pp. 645-667, https://www.jstor.org/stable/25046265.

Shapiro, C. and H.R. Varian (1998), *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press, Boston.

Strubell, E., A. Ganesh and A. McCallum (2019), "Energy and policy considerations for deep learning in NLP", *arXiv*, preprint arXiv:1906.02243, https://doi.org/10.48550/arXiv.1906.02243.

Utterback, J.M. and W.J. Abernathy (1975), "A dynamic model of process and product innovation", *Omega*, Vol. 3/6, pp. 639-656, https://doi.org/10.1016/0305-0483(75)90068-7.

Whittaker, M. (2021), "The steep cost of capture", *Interactions*, Vol. 28/6, pp. 50-55.

Wooldridge, M. (2020), *The Road to Conscious Machines: The Story of AI*, Penguin, London.

Notes

[1] While it is likely that education will become increasingly "datafied", concerns about privacy and the challenges of measuring educational outcomes will tend to hinder the deployment of AI systems at the scale seen in the web and media sectors.

[2] Arguably, sufficient generalisability in a single (dominant) AI design could make it repurposable for all use-cases. However, such generalisability still seems far away, again providing reasons for preserving research diversity around such topics as AI techniques suitable for low-data, fast-changing and high-stakes contexts where deep learning techniques currently under-perform.

[3] To identify AI papers, salient terms for papers have been extracted that have been classified by their authors in machine learning/neural network categories. Papers with high frequency of those terms have been sought outside of those categories.

From:

Artificial Intelligence in Science

Challenges, Opportunities and the Future of Research

Access the complete publication at:

https://doi.org/10.1787/a8d820bd-en

Please cite this chapter as:

Mateos-Garcia , Juan and Joel Klinger (2023), “Is there a narrowing of AI research?”, in OECD, *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/77709ef0-en