R SHEET ON
COMPETITION IN ARTIFICIAL INTELLIGENCE

November 2018

BACKGROUND

Artificial Intelligence (AI) is developing rapidly, and countries from around the globe are beginning to articulate national strategies for handling the political ramifications. Powering innovations like driverless cars, autonomous drones, full-sequence genetic analytics and powerful voice-assistant technology, the future certainly looks full of potential. However, unsettled questions about who will reap these benefits and when they will be achieved leave storm clouds on the political horizon.

Formal definitions for AI vary but generally the term can be used to refer to the broad suite of computer algorithms being used to automate or improve aspects of human decision making. In the most current iteration, this is largely being accomplished via machine learning (ML), whereby an algorithm uses statistical techniques to find patterns from a dataset and progressively improve prediction ability at a given task (email spam filters are a great example). In this understanding, AI exists on a spectrum rather than a binary, with increasing sophistication in the ability to apply various models to solve the problem at hand, indicating higher levels of intelligence.

CURRENT DEBATE

Amid concerns about the degree of industrial concentration and about the slow diffusion of AI applications across the rest of the economy, there is an underexamined overlap that connects these two concerns to the same set of policies: namely, high barriers to entry due to supply-side constraints.

There are significant barriers to entry in AI development and application, many of which are the direct result of government policies. These barriers have inadvertently boosted the market power of incumbent firms and by reducing them, we may enable new firms to better compete, while also removing some of the bottlenecks that slow down research and integration of AI systems across the entire economy.

SUMMARY

- AI can power innovations across the entire economy by augmenting human decision-making, automating rote tasks and finding new patterns in datasets.

- To ensure a competitive ecosystem, Congress should focus on reducing barriers to entry that exist in the supply of labor, the supply of data and access to hardware.

- Congress can increase the supply of skilled AI analysts by reforming our immigration system to allow more high-skill AI talent, and by allowing companies to deduct the cost of training AI talent.

- Congress can increase the supply of high-quality datasets by opening up more government data for public use, and by clarifying the fair-use exemption for training data.

- Congress can help ensure access to specialized AI hardware by avoiding political destabilization of international supply chains, and by maintaining our healthy ecosystem around distributed platforms.

SUPPLY OF TALENT

Perhaps the single biggest bottleneck in AI development and application today is the supply of skilled data scientists and machine-learning engineers. Typical AI specialists can expect to earn between $300,000 and $500,000 per year at top tech firms—numbers that are significantly higher than their peers in other computer-science-related subfields.

If there were appropriate policy levers to increase the supply of skilled technical workers available in the United States, it would disproportionately benefit smaller companies and startups, making the overall ecosystem more competitive, while simultaneously increasing the rate of AI diffusion in other industries.

To begin alleviating this labor shortage, Congress should first reform our immigration system to allow more high-skilled international students to stay in the country upon completion of their degree. In 2015, the United States had

58,000 graduate students in computer science fields, the overwhelming majority of which (79 percent) were international. This represents a significant portion of the overall AI talent supply being cultivated each year, as students from all over the world are attracted to the nation's higher education system.

Secondly, Congress should allow companies to deduct the expense of training AI workers from their taxes in the same way they deduct capital investments. The current business expense tax deduction limits the scope of worker training to increasing productivity only in jobs the worker already occupies. Expanding this tax deduction in size and scope would increase the incentive for employers to invest in AI training, and it would equalize employer incentives between capital and labor investment.

SUPPLY OF DATA

To expand the supply of high-quality datasets needed to train new ML models, policymakers should first examine the datasets they already own that could be made public. For example, many cities and municipalities have useful data around traffic patterns, electricity usage and business development that, if opened, could lead to reduced-cost service provision and better analytics. And, there have been a flurry of recent pushes in Congress to standardize the publication of government-agency datasets in a machine-readable format.

There is also the matter of industries in which open data might become the norm if existing regulations are relaxed or streamlined. The healthcare industry seems like a particularly promising target in this respect, as HIPAA has long been considered a barrier to the development of data sharing across medical professionals and companies. Allowing consumer health data to be more easily shared, with the proper privacy safeguards, could enable a renaissance in drug development and personalized medicine, as recent ML advances have proved quite promising when appropriate data have been made available.

In addition to making more government datasets opensource, we should also take a second look at some of the intellectual-property laws that intersect and interact with the ML process.

It is currently ambiguous whether or not using copyrighted works as training data in an ML model would legally qualify under the fair-use exemption. This legal uncertainty is a barrier to publicly using a large number of copyrighted works as training data. This disproportionally hurts startups that do not have access to the reams of consumer data that large incumbents have. Congress should further study and clarify the fair- use exemption to create new opportunities for startups to compete.

ACCESS TO SPECIALIZED HARDWARE

As the hardware used to train ML models becomes more specialized, the international supply chains that support the distribution of this hardware become even more important. Large incumbents are able to strategically onshore the production of particular AI chipsets that they know are politically vulnerable to trade shocks. Startups do not have this same ability. Accordingly, the stability of international supply chains becomes a crucial issue for ensuring broad access to specialized AI chips.

**Note: References and a deeper discussion of the issue are available in the full-length study online at: https://www.rstreet.org/2018/10/09/reducing-entry-barriers-in-the-development-and-application-of-ai.

CONTACT US

For more information on this subject, contact the R Street Institute: 1212 New York Ave NW Suite 900, Washington D.C. 20005, 202.525.5717

Caleb Watney
Technology Policy Fellow
cwatney@rstreet.org
620-204-1127
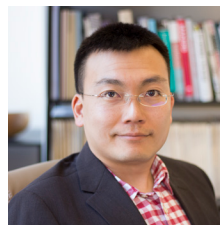
Charles Duan
Director of Technology and Innovation Policy
cduan@rstreet.org
202-900-8247