Berkeley
UNIVERSITY OF CALIFORNIA

BERKELEY ROUNDTABLE ON
THE INTERNATIONAL ECONOMY

BRIE / WITS
Technology Briefing #2

October 2019

AI and Domain Knowledge: Implications of the Limits of Statistical Inference

Drafted by **Christopher Eldred**
Based on presentations by **Michael Borrus** and **Alberto Sangiovanni-Vincentelli**
Significant editorial contributions from **John Zysman** and **Mark Nitzberg**

Acknowledgements: This research was funded in part by the Ewing Marion Kauffman Foundation, and Denk Fabrik, the Policy Lab of the German Federal Ministry of Labour and Social Affairs. The contents of this publication are solely the responsibility of the authors. Others made contributions, including Jane Macfarlane and Matthew Stepka.

AI and Domain Knowledge: Implications of the Limits of Statistical Inference

Introduction

"With artificial intelligence, we are not crawling or walking or running," booms the voice of actor and musician Common in 2018's Microsoft AI TV spot. "We are flying."[1] This ad probably resonated with casual audiences who saw it throughout the NBA playoffs that year and beyond. More than any other aspect of digital technology, AI inspires imagination. Whether optimists or pessimists about AI's impact, most see it as very powerful, becoming more powerful still, and perhaps limitless in the possibilities of what it can do.

AI, which today principally relies on a discipline of computer science called machine learning (ML), is indeed powerful. Recent advances in the use of ML have powered an AI system that is incomprehensibly skilled at Go, a game once thought to be far too complex for anyone without a human's strategic intuition.[2] AI-based tools can now surpass humans in the reliable recognition of images and the reliable transcription of voices.[3] Decent-to-good language translations are instantly available to anyone with an internet connection. These capabilities have an extraordinary number of applications with the potential to transform economies and societies.

But ML-powered AI does have fundamental limitations. And these limitations have implications for AI's development, application, and impact in the coming years. This paper will first describe ML's fundamental nature as a form of statistical inference, and explore how this constricts effective use of today's AI to certain kinds of problem domains. Through the lens of this limitation, it will then take a position in three debates over how AI will unfold:

- First, will AI's diffusion throughout the economy and society be led by firms with the most advanced AI technical capabilities? Or will expertise in the problem domains in which AI tools are applied, rather than the technical sophistication of the tools themselves, be the more important determinant of AI's effectiveness and market success?
- Second, will China's access to unmatchable volumes of data lead to a dominant position in global markets for AI tools? Or will the importance of access to local data around the world place Chinese firms on equal footing with domestic companies or other global rivals?
- And third, is the creation of AI that can match or exceed human intelligence, or Artificial General Intelligence (AGI), a foreseeable circumstance? Or is it a possibility only in the distant future, if it is possible at all?

We believe that the absolute importance of narrow domain knowledge in effective use of AI biases markets in favor of companies with expertise in applying AI to narrow problems over technical behemoths, deflates the hype that lifts up China's potential as a hegemonic AI superpower, and makes AGI a very remote possibility indeed.

Through all three of these debates runs a question of the importance to AI of broad reach and massive scale. Keeping in mind the sudden rise to unprecedented size of network effect-aided platforms, it is easy to assume that deep networks running on huge computers analyzing datasets that stretch beyond the horizon will likewise dominate markets, determine global economic leadership, and possibly take over the world. We make the case that while its impact will doubtless be profound, the vehicles of its future will be more modest in scope. Instead of worrying about the impact of AI at massive scale, policymakers, investors, and other stakeholders would do well to focus on deftly applying AI to the unique and varied problems throughout business and society.

The Limits of Statistics

The term “AI” has appeared suddenly and in abundance in recent years, but AI has existed for a long time and taken many forms along the way. As mentioned above, today’s most widely applied AI systems rely on machine learning (ML). Like any algorithms, ML algorithms produce outputs based on the input of information provided by the environment or user. Every ML system chooses which output to produce based on sets of example output/input pairs that they have seen before. It is through analyzing the relationships between inputs and outputs – or “training” – that ML systems “learn” which output has the highest probability of correctly matching a given input.

ML algorithms, whether they transcribe speech or identify animals in photographs, are little more than reflections of the correlations between inputs and outputs in their training data. In the former, the system knows which letters arranged in a particular order match the sounds a human voice is making because it has seen the same or very similar pairings of letters and sounds in its training data. In the latter, the system produces the word “dog” when presented with a photograph whose pixels constitute an image of a canine because it has been trained on a dataset that includes many photographs with similar pixel arrangements paired with the word “dog.” This kind of analysis, where ML algorithms infer what to do based on a statistical analysis of data, powers most of the headline-writing AI achievements and “AI-powered” software features that have become commonplace.

ML cannot handle inputs that are statistically eccentric compared to those examples they have seen in their training data; put another way, they will not know what to do in situations they have never seen before. This means that today’s AI systems can only work reliably well in decision-making domains with two features. First, they must be non-dynamic, in that they must not frequently present an AI system with situations or information that it hasn’t seen before. Second, they must be narrow, meaning that the universe of possible inputs and outputs must be able to be comprehensively described by training data.

Put another way, the context of analysis must be rigidly defined and maintained. But in our fluid social and economic world, contexts are ever changing and evolving. This does not mean that AI is not useful, but helps define the boundaries of its usefulness.

Progress in AI over the last decade has stemmed largely from two complementary forces. First, the spread of mobile technology led to the creation of exponentially increasing amounts of text, recordings, images, and other data for AI to statistically analyze. Second, the increase of computing power described by Moore’s Law, in which processing speeds have doubled approximately every two years going back four decades, allows algorithms to perform this analysis quickly and at scale.[4] It’s these two forces together that have made AI particularly adept at voice transcription, image recognition, translation, and game playing.[5]

The collection of more kinds of data expands the domains that AI can work in, and evolutions in technique and computing power continue to help AI make better decisions faster. But no ML algorithm so far can overcome the need for narrow and relatively static domains that statistical inference requires. AI today is used exclusively in non-dynamic, narrow problem domains; the most advanced AI techniques simply do not allow for applications that can generalize from one problem domain to an entirely different one.

Does the importance of narrow domain knowledge, combined with an understanding of how to use effectively AI tools, give companies with expertise in problem areas to which AI is applied a market advantage over other firms, regardless of their technical sophistication?

In the pursuit of boundary-pushing innovation and transformative applications in AI, many firms place a premium on technical prowess. Firms like Google, Facebook, Amazon, and Microsoft have lured many AI scholars to their in-house research teams with generous compensation packages, hoping that the brightest theoretical minds will build them the most sophisticated AI algorithms.[6] And many outside observers assume that such firms have the right approach; surely, the companies with the most advanced deep learning models and computing power will lead the diffusion of AI throughout the economy and society. Just in September, an article headline in *Fast Company* implied that Google's computing power and AI talent position it to build machine systems that approximate human intelligence in foreseeable future.[7]

Obviously, using the best algorithms and most computing power is no handicap for an AI application. And in some areas, having the top hardware and software may indeed determine which application is most effective. This could be especially true for tools addressing problem domains that are very similar across diverse firms and industries, for which assembling an output-comprehensive dataset is relatively easy. Consider AI software that captures handwritten information from paper forms and transforms it into digital data. Myriad industries have to process large numbers of forms filled out by consumers and employees. Most forms follow fairly similar formats and structures. People's handwriting on forms typically falls within a reasonably predictable range of styles. It makes sense that a solution built on the most powerful algorithms and computers are the fastest and most accurate at producing the right digital information when presented with scribbles on paper from a doctors' office. And it is reasonable to expect that the firms with state-of-the-art technical prowess would be able to assemble an output-comprehensive dataset for such a typical problem.

But many problems that AI can help solve are unique to a particular industry or function in a workplace. And the critical underpinning of firms most successful in using AI to solve these kinds of problems is expertise in the relevant problem and how to use AI to solve it, rather than technical robustness or sophistication. Indeed, some firms who may not have invested in top AI talent still have a better idea of the niche domains within the huge variety of business problems in the economy. They are also more likely to know what data is required to create the output-comprehensive dataset needed to solve them. With an output comprehensive dataset, these firms do not need the most cutting-edge machine learning model to delivery satisfactory solutions. Meanwhile, the firms with the best AI hardware and software may not even know the problem exists, let alone know how to begin using statistical inference to solve it.

This is why a wealth of firms are emerging that offer focused AI solutions to narrowly defined problems. One example is Biota, a firm that helps energy companies determine the size, interconnectivity, and other valuable information about shale gas or oil deposits, with minimal drilling.[8] They do this by sampling DNA from microorganisms living in the hydrocarbons and running them against their massive dataset of information about other deposits around the country.[9] Only a firm with deep expertise in the particular problems of the energy industry would have been able to design and build this solution. Another example is Lilt, a firm that specializes in combining AI with human translators to achieve consistent, high-quality translations for professional contexts.[10] Such firms are often born of collaborations between people with technical talent and people with direct experience in the industry problems the solutions are being developed to solve.

Theoretically, it's not impossible for firms with the highest technical ability to also develop expertise in niche problem domains and build solutions to address them. Look at the AI offerings of companies like IBM and Microsoft, and one will find solutions for specific industries from health to education to manufacturing.[11] But most of the time, firms who can afford to hire the most in-demand AI talent and purchase the largest, most advanced computers are going to be such large firms with economy-wide reach; in practice, these firms are simply unlikely to develop the narrow, domain-comprehensive datasets needed to solve a multitude of narrow business problems. The most successful AI applications developed by these firms will likely be limited to the broadest possible problem areas. Meanwhile, smaller firms will more naturally arrive at the domain expertise needed to leverage AI everywhere it can be applied.

This circumstance has bearing on the decisions made by policymakers globally who are looking to develop a competitive advantage for domestic technology firms and their AI applications. In particular, the European Union recently launched the AI4EU project, which seeks to pool together the knowledge, algorithms, tools, and resources of over seventy corporate and academic partners to develop a comprehensive On-Demand AI Platform.[12] In doing so, the EU may be seeking to match what they perceive to be the competitive strengths in AI of American tech giants, hoping to forestall in this emerging field the kind of dominance they achieved in cloud computing, search, social media, and other dimensions of digital disruption. But to the extent that narrow domain expertise is the most important factor in creating a competitive advantage in AI, European resources for AI development would be better directed to alternate strategies.

Does the importance of domain-specific knowledge and AI understanding suggest that China will not necessarily become the globally dominant force in AI, despite unrivaled data scale? Will not international AI competition rather be determined by the relative skill in applying AI tools?

Just as it is common to assume that the largest firms will lead the way in AI, so is it common to hear the same about the largest country. China's central government has made leadership in AI a strategic priority, and some observers believe China is on its way to AI dominance.[13] Many believe that China is on its way to dominance in AI, and that this will be a pivotal factor in China's rise to economic supremacy. These analyses typically point to China's data advantage: with a population of 1.4 billion, pervasive deployment of CCTV cameras, cutting-edge adoption of mobile payments, peer-to-peer lending, and other technologies, the nation has more data than any other country by far.[14] Meanwhile, the Chinese government plays an active role in sharing data from its unmatched surveillance apparatus with select AI companies.[15] With all this data, will not Chinese companies build the most robust, output-comprehensive datasets and train the most reliable and accurate AI systems that dominate domestic competition in markets around the world?

On the contrary, the overwhelming volume of Chinese data does not always make it best equipped to build AI solutions that work well in other countries. As noted above, the sheer volume of data is not the only factor in training the highest-performing AI algorithms; it is most important that training data accurately and comprehensively represent the universe of possible input/output pairs the system will have to contend with. And in many cases, data from China will not be comprehensive in describing inputs and outputs in other countries, even in similar business contexts.

To be sure, there are some domains in which Chinese data may prove sufficient to train successful AI systems in other countries. These are domains that are very similar regardless of country. As a thought experiment, consider cancer diagnosis: one imagines that cancer is likely to have the same visual signature in medical imaging regardless of whether the patient is in Shanghai, Seville, or Santiago. If this is true, a dataset of medical imaging and diagnosis information from China's massive patient population might comprehensively describe all possible pairings of diagnoses and image patterns for patients in Germany or Nigeria. Another domain that is similar regardless of locale could be the counting of people in a public square under video surveillance: people look similar enough at a distance regardless of locale for a system trained to count people in a square in one country to do so in any other. China may be able to build superior applications addressing these domains than domestic companies in other countries can manage to create.

But many other domains are likely to be too dissimilar across borders for applications trained on Chinese data to work well in another country. Consider retail apparel: a Chinese company's recommendation engine might be the highest-performing in the world at successfully matching its customers with clothing styles they would like to buy. But tastes and preferences in clothing vary from country to country; an algorithm trained on Chinese data might be able to predict that German males generally prefer to purchase trousers over dresses, but might not be able to make recommendations as to the most popular and stylish trousers among German consumers that were not a part of its training sample. Domains where cultural preferences are a relevant force will make it particularly difficult for algorithms trained in one country to work well across the globe.

This is not to say that China has no advantage. Some domains are not identical from country to country, but bear important fundamental similarities. In these cases, China may be able to build AI systems that are effective in other countries through transfer learning, an AI technique wherein an algorithm trained to perform one task can be repurposed to perform a second task with fundamental similarities to the first task using a relatively small amount of new training data.[16] This could potentially be useful in creating an AI system to automatically diagnose diseases; while some diseases are much more common in certain countries than in others, the vast majority of patients using this AI system would likely have symptoms that are common globally. Conceivably, an algorithm could be trained to recognize these common conditions better than any other by analyzing China's massive population; this algorithm could then be fine-tuned by feeding it training data describing patient conditions particular to a new country. It may be easier to adapt an AI system for medical diagnoses from China via transfer learning than it would be for each country to build its own such system from scratch.

Other factors are important in considering whether China will have a competitive advantage in developing AI applications for export. For applications where transfer learning is appropriate, China will still have to acquire the relevant fine-tuning data. Additionally, even in domains where there are major differences between China and other countries, domestic firms may struggle to find the financing and talent they need to develop their own solution. But at a minimum, the need for domain-specific data places China on equal footing in building globally-competitive AI tools with the US and other nations with robust AI industries. Overall, China's advantage in developing AI tools for its domestic market conferred by its mountain of data by no means equates to an advantage in global markets.

Narrow, effectively deployed tools will define AI's path for the foreseeable future. Artificial general intelligence will not be created through continuous progression in machine learning.

Yet while recent progress is enormous relative to past progress, it is tiny compared to the overall journey towards AGI. Some experts view human intelligence as fundamentally separate from AI in ways that will not be bridged anytime soon. Fundamental to this distinction is every human's abstract model of the world that we are constantly referring to and updating as we go about our lives.[17] In any situation, this abstract model gives us a subconscious idea about what is happening and why.[18] This, in turn, informs common sense, which enables us to handle situations that we haven't seen before.[19] It's this foundation that allows us to act intelligently across an endless number of uncertain and complex situations, from carrying a cup of coffee through a crowded office kitchen to determining the best strategy to help a business grow.

This abstract model of the world stems from things that are difficult or impossible to artificially replicate: our lived experience, and the millions of years of evolution that shape our instinct and intuition. Few would suppose that knowledge accumulated over years of living life is not essential to making decisions and handling situations. And some things that we know how to do by instinct, such as grasping a bottle of honey or reaching down to break a fall, are our heritage from natural selection.

Certain AI applications may be nearly impossible to perfect without revolutionary breakthroughs in computer science. Such applications could include autonomous driving; while it has been widely assumed by many that fully autonomous vehicles will become widespread in just a few years, vehicles will always encounter situations on busy streets that they have not seen before; in such cases, human intervention will be required.[20]

Taking this perspective into account, it may be a mistake to view human and artificial intelligence as if they exist on the same continuum. Certainly, as the amount and variety of data collected continues to expand and computing power grows, AI systems will be able to perform feats of analysis and decision making that humans could never match.

Rather than directly comparing human intelligence to AI, it may make more sense to explore how they can capitalize on one another's strengths to achieve better outcomes. The benefits of the "complementarity" between human and machine intelligence will be explored in greater detail in a later work.

Conclusion

It is not hard to argue that in recent years, AI has entered a new phase. The most advanced AI algorithms are so complex that we cannot fully explain their decisions.[21] This lends them an air of agency that we tend to only associate with other living things. For the first time, we can tell what the computer is saying, but we can't really tell what it exactly is thinking. And the statistical inference techniques of today's AI are indeed powerful. But they have fundamental limits.

Understanding of the problem domains that AI is meant to solve remains the central determinant of AI's success, not algorithmic sophistication or other technical prowess. The greatest positive impact of AI will come from those with expertise in learning *how* to apply AI methods, more than from the development of more sophisticated AI techniques; meanwhile, the largest number of opportunities for AI investment lie in discovering and nurturing applications in these narrow problem areas, rather than building the most powerful and technically sophisticated systems. In addition, massive piles of data do not necessarily give China a significant advantage over other global powers, or even over domestic technology firms in other countries across the world, in developing AI solutions; rather, skill at developing applications tailored to local problems will be the determining factor in AI's global leadership. Finally, the worst fears about AGI are exaggerated; machines that have the flexibility of human intelligence are not on the horizon. We have more to fear from quantum computing than we do the possibility of an ML-based system surpassing humans anytime soon.

This paradigm may yet shift. Research is underway on techniques that can generate their own comprehensive sets of training data, which could power applications that are effective in domains where little data exists.[22] Still other areas of research are pursuing strategies wherein programs analyze and evaluate human behavior in different situations to model the best course of action.[23]

But all of these avenues of research are in their early phases. Taking into account the recent history of AI, and the logic underlying many recent breakthroughs, many of the presupposed impacts of ML-powered AI are unlikely to transpire anytime soon. By understanding the limits of statistical inference on AI, stakeholders can be better equipped to effectively leverage this powerful tool in economies, societies, and the world at large.

[1] "Microsoft AI Commercial: Empowering Innovators ft. Common," April 22, 2018, *Microsoft Youtube Channel,* accessed October 14, 2019, https://youtu.be/Z5OWdqfAYfw?t=6.

[2] Steven Borowiec, "AlphaGo seals 4-1 victory over Go grandmaster Lee Sedol," March 15, 2016, *The Guardian,* accessed October 14, 2019, https://www.theguardian.com/technology/2016/mar/15/googles-alphago-seals-4-1-victory-over-grandmaster-lee-sedol.

[3] "AI Progress Measurement," *Electronic Frontier Foundation,* accessed October 14, 2019, https://www.eff.org/ai/metrics#Speech-recognition; https://www.eff.org/ai/metrics#Vision.

[4] Eric Martin, "Moore's Law is Alive and Well," December 21, 2018, *Medium Predict,* accessed October 14, 2019, https://medium.com/predict/moores-law-is-alive-and-well-adc010ea7a63.

[5] "AI Progress Measurement," *Electronic Frontier Foundation,* accessed October 14, 2019, https://www.eff.org/ai/metrics#Speech-recognition; https://www.eff.org/ai/metrics#Vision; https://www.eff.org/ai/metrics#Translation; https://www.eff.org/ai/metrics#Game-Playing.

[6] Tracy Malingo, "The Misguided Rush of the Academic AI Brain Drain," August 23, 2018, *The New Stack,* accessed October 14, 2019, https://thenewstack.io/the-misguided-rush-of-the-academic-ai-brain-drain/.

[7] Katrina Brooker, "Google's quantum bet on the future of AI – and what it means for humanity," September 17, 2019, *Wired,* accessed October 14, 2019, https://www.fastcompany.com/90396213/google-quantum-supremacy-future-ai-humanity.

[8] "Biota Technology," *Biota Technology,* accessed October 14, 2019, http://www.biota.com/.

[9] "Biota: Revealing the Reservoir," April 19, 2016, *Biota Technology's Youtube Channel,* accessed October 14, 2019, https://www.youtube.com/watch?v=X9rWYQQSYDo.

[10] "About," *Lilt,* accessed October 14, 2019, https://lilt.com/about.

[11] "Enterprise-ready AI," *IBM,* accessed October 14, 2019, https://www.ibm.com/watson/about; "Intelligent applications," *Microsoft,* accessed October 14, 2019, https://www.microsoft.com/en-us/ai/intelligent-apps.

[12] " Welcome to the AI4EU Platform," *AI4EU,* accessed October 14, 2019, https://www.ai4eu.eu/.

[13] Gregory C. Allen, "Understanding China's AI Strategy," February 6, 2019, *Center for a New American Security,* accessed October 14, 2019, https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy; "China is Winning the AI Race: Kai-Fu Lee Explains | Thrive Global," September 21, 2018, *Thrive Global Youtube Channel,* accessed October 14, 2019, https://www.youtube.com/watch?v=IMit5cPXbTo.

[14] Saheli Roy Choudhury, “As information increasingly drives economies, China is set to overtake the US in race for data,” February 13, 2019, *CNBC,* accessed October 14, 2019, https://www.cnbc.com/2019/02/14/china-will-create-more-data-than-the-us-by-2025-idc-report.html; Yu Yifan, “China’s AI industry has this one huge advantage over the US,” July 23, 2017, *South China Morning Post,* accessed October 14, 2019, https://www.scmp.com/business/article/2103791/chinas-ai-industry-has-one-huge-advantage-over-us.

[15] Josh Horwitz, “The billion-dollar, Alibaba-backed AI company that’s quietly watching people in China,” April 16, 2018, *Quartz,* accessed October 14, 2019, https://qz.com/1248493/sensetime-the-billion-dollar-alibaba-backed-ai-company-thats-quietly-watching-everyone-in-china/.

[16] Scott Martin, “What Is Transfer Learning?” February 7, 2019, *Nvidia,* accessed October 14, 2019, https://blogs.nvidia.com/blog/2019/02/07/what-is-transfer-learning/.

[17] “How smart is today’s artificial intelligence?” December 19, 2017, *Vox Youtube Channel,* accessed October 14, 2019, https://youtu.be/IJKjMIU55pE?t=364.

[18] Clive Thompson, “How to Teach Artificial Intelligence Some Common Sense,” November 13, 2018, *Wired,* accessed October 14, 2019, https://www.wired.com/story/how-to-teach-artificial-intelligence-common-sense/.

[19] Clive Thompson, “How to Teach Artificial Intelligence Some Common Sense,” November 13, 2018, *Wired,* accessed October 14, 2019, https://www.wired.com/story/how-to-teach-artificial-intelligence-common-sense/.

[20] Nidhi Kalra, Susan M. Paddock, “Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?” 2016, *Rand Corporation,* accessed October 14, 2019, https://www.rand.org/pubs/research_reports/RR1478.html.

[21] “A Brief History of Machine Learning Models Explainability,” September 24, 2018, *Zelros AI Medium,* accessed October 14, 2019, https://medium.com/@Zelros/a-brief-history-of-machine-learning-models-explainability-f1c3301be9dc.

[22] Jack Clark, “Import AI 145,” May 6, 2019, *Import AI,* accessed October 14, 2019, https://us13.campaign-archive.com/?u=67bd06787e84d73db24fb0aa5&id=9f4b66162f&e=fcc7585387; Qizhe Xie, Zihang Dai, Eduard Hovy, Minh-Thang Luong, Quoc V. Le, “Unsupervised Data Augmentation for Consistency Training,” April 29, 2019, *arXiv,* accessed October 14, 2019, https://arxiv.org/abs/1904.12848; Rui Wang, Joel Lehman, Jeff Clune, Kenneth O. Stanley, “Paired Open-Ended Trailblazer (POET): Endlessly Generating Increasingly Complex and Diverse Learning Environments and Their Solutions,” January 7, 2019, *arXiv,* accessed October 14, 2019, https://arxiv.org/abs/1901.01753.

[23] Rohin Shah, Dmitrii Krasheninnikov, Jordan Alexander, Pieter Abbeel, Anca Dragan, “Preferences Implicit in the State of the World,” February 12, 2019, *arXiv,* accessed October 14, 2019, https://arxiv.org/abs/1902.04198; Anca D. Dragan, “Robot Planning

with Mathematical Models of Human State and Action," May 11, 2017, *arXiv,* accessed October 14, 2019, https://arxiv.org/abs/1705.04226.