Export Control Reform Implementation: Outside Perspectives

Senate Banking Committee

Prepared Testimony and Statement for the Record of
Ben Buchanan

Assistant Teaching Professor, School of Foreign Service
Senior Faculty Fellow, Center for Security and Emerging Technology
Georgetown University

Thank you, Chairman Crapo and Ranking Member Brown, for holding this important hearing and for inviting me to testify.

My name is Ben Buchanan. I am an Assistant Teaching Professor at the School of Foreign Service and a Senior Faculty Fellow at the Center for Security and Emerging Technology, both at Georgetown University. I am also a Global Fellow at the Woodrow Wilson International Center for Scholars, where I teach introductory classes on Artificial Intelligence and cybersecurity for congressional staff. My research specialty is examining how cybersecurity and AI shape international security. I co-authored a paper entitled "Machine Learning for Policymakers."[1]

As this committee is well aware, export controls are legal tools that are applied to technology. If either the tool or the technology is not a good fit, export controls will fail. Given the expertise of my two fellow witnesses on the legal nuances of the tools themselves, I believe I will be of most value to the committee in talking about some of the technologies in play and what makes export controls comparatively more or less suitable with these technologies. As a way of opening our discussion, I will focus on one particular suite of technologies that is particularly notable, artificial intelligence, but I believe this discussion will also apply to other relevant technologies.

Conceptualizing AI

Nobody has a crystal ball, but there are other ways to consider our modern and near-future era of AI that will be useful for this discussion. To do so, it is important to understand how AI differs from so much of what came before it. An analogy will help.

One can imagine two ways of teaching a child to perform a task. The first is to give very clear instructions in a language the child understands about what the task is and how it is to be performed. The second is to show the child, through a series of examples, how the task works, and have the child infer the important rules and patterns necessary to get the job done. At various points in a child's education, they learn different tasks through each of these methods.

Traditional software development, and even some older versions of AI, work in a way that is similar to the first method. They rely on software developers understanding the problem to be solved in great depth, and then imparting this expertise to the system. For

[1] Buchanan, Ben and Taylor Miller. "Machine Learning for Policymakers." *Belfer Center for Science and International Affairs* (2017), https://www.belfercenter.org/sites/default/files/files/publication/MachineLearningforPolicymakers.pdf.

example, in a program designed to play chess, the software developers may consult with grandmasters to understand the optimal strategies for a wide range of situations, and then program those ideas into the code.

Modern AI systems, known as machine learning systems, use the second method, the one involving inference. In a machine learning system, rather than receive clear instructions about how to do the task, software developers create an algorithm that determines how the system should learn. They then provide that algorithm with lots of relevant data and computational power (the processing hardware that makes machine learning algorithms function).

There are thus three parts to this system: the algorithm, the data, and the computational power. Together, they form an essential triad. Each is more or less important in various versions of machine learning, but at the same time, each in its own way is critical. To understand why, it is worth examining the triad in a little more detail.

Data

It is in vogue to say that data is the new oil. This is because, to use the second kind of program I described above--the machine learning method--a lot of relevant data is often required. From this data the machine learning system will infer important patterns and nuances, and will determine what success and failure look like. It is thus vital that the data provided to the machine be representative of the problem in all its complexity and plentiful.

A large part of the reason that companies like Google, Amazon, and Facebook are successful with the AI systems they deploy is because they aggregate gigantic amounts of data. In essence, the large datasets these companies assemble provide them with a competitive advantage over others. Large companies based in other nations, such as China's Baidu, Alibaba, and Tencent, derive similar advantages from their datasets. It seems to me that export controls are unlikely to be of much use in managing this competition or guarding against potential threats from data, both because companies already have an incentive and tools to secure and not share their assembled data and because export controls are comparatively ill-equipped to stop the transfer of sensitive data relative to other tools like classification (for government data), and licensing or contractual restrictions regardless of export.

Algorithms

Algorithms are the second component of the AI triad. These software instructions dictate *how* the machine learning system will learn. They stipulate how it will interpret the data,

what sort of capabilities it will develop, and what inferences it will learn to draw that can be applied to future tasks. There are a wide variety of algorithms, each suited to different kinds of tasks, from classifying images to making predictions about housing prices based on historical trends, to generating new pictures of people who look real but do not actually exist. The algorithmic frontier is rich, and a great deal of progress has been made in the last seven years.

The prevailing ethos is that, once an algorithmic advance is made, researchers post it online and share it with others. In this sense, AI research is remarkably open, far more so than the fierce competition of the technology industry would normally suggest. There are exceptions to this practice, instances in which algorithms have not been published due to national security concerns--most notably a decision by OpenAI, a leading research lab, not to publish a powerful algorithm that could be used to generate realistic-fake text.

That said, the experience of several decades has shown that government efforts to control the export of computer code are usually futile, and I think it is fair to say that export controls are unlikely to be useful in stopping all but the most powerful of algorithms. And even with those most powerful algorithms, I have doubts about the suitability of our current list-based export control systems, given the changing pace of technology and the movement of the technological frontier.

Computing Power

This brings us to the last part of the triad: computing power, or what AI researchers simply call “compute.” It is easy to ignore, but it remains vitally important, perhaps prohibitively so. In the last seven years, we have witnessed a revolution in computing power applied to machine learning. One study by OpenAI indicated that between 2012 and 2018, the computing power applied to top machine learning systems increased by a factor of 300,000; if a cell phone battery lasted one day in 2012 and increased by the same factor, that battery would now last 800 years.[2]

There is much to discuss about why this increase in computing power has occurred, but the most salient factor for our purposes today is that, unlike algorithms and data, computing power is a function of hardware, not software. That is, computers are tangible products that are easier to manage, including with export controls. My judgment is that, to the degree that export controls are relevant to the problem of managing AI and other technologies such as 5G, it will controls on this hardware component, and likely on the

[2] “AI and Compute”, OpenAI, (2018), https://openai.com/blog/ai-and-compute/

hardware that manufactures specialized computer chips for AI. This statement is both a commentary on the limitations of export controls to the problem but also on the more narrow areas where they might be suitable for protecting national security.

To be clear, in order for any such controls to work--whether on AI hardware or something else--they must be conducted in a multilateral fashion with allies, given that a great deal of hardware engineering expertise is outside the United States.

I thank you again for holding this hearing and the opportunity to lay out the basics of this complicated, fast-changing field for your consideration as you review the implications of export control for AI and other technologies. As you know, it is vital that we both protect national security and not squash innovation. This is an area that the Center for Security and Emerging Technology has been studying, and we expect to publish our analysis on it in the weeks to come. In the meantime, I look forward to your questions.