TESTIMONY OF SAMM SACKS
Hearing on Promoting Competition, Growth, and Privacy Protection in the Technology Sector
Senate Finance Subcommittee on Fiscal Responsibility and Economic Growth
December 7, 2021

Chair Warren, Ranking Member Cassidy, and Members of the Subcommittee, thank you for the opportunity to testify today.

I am a Senior Fellow at Yale Law School's Paul Tsai China Center and a Cybersecurity Policy Fellow at New America. I have worked as an analyst of Chinese data and technology policies for the last decade, in the U.S. national security community, and in the private sector. I also advise corporate clients on China's technology policies.

Today I will focus my testimony on data security in the context of the U.S.-China relationship and global cross-border data flows.

While my expertise focuses on China—and I will first speak specifically about the Chinese government's approach to acquiring and extracting value from data—my view is that the most effective solutions for the United States require a more comprehensive approach to regulating data security and privacy. Some of these challenges require tools that are specific to risks posed by China, but these issues are bigger than China. Setting basic standards on what data can be collected and retained by all companies will help protect U.S. personal and other sensitive data, regardless of whether the risk comes from a state-sponsored hacker, a data broker, or a private company transferring the data to China. U.S. lawmakers have an opportunity to address transnational security threats while also advancing a more secure, ethical, and democratic global internet in its own right.

China's National Data Strategy

1. *The Chinese government has embarked on an ambitious national data strategy with the goal of acquiring, controlling, and extracting value from large volumes of data.*

In addition to China's two landmark laws that took effect this fall (the Data Security Law and Personal Information Protection Law[1]), Beijing has elevated the concept of data as an economic and strategic asset[2], centralizing state power over information flows within and outside of China's borders:

[1] For translation and analysis of the Data Security Law, Personal Information Protection Law, and related regulations and directives, please see the Stanford Cyber Policy Center's DigiChina Project, https://digichina.stanford.edu/.
[2] The concept of data as a strategic resource is not new in China. It appears in the Big Data White Papers (2014, 2016, 2018) published by an influential think tank under the Ministry of Industry Information Technology (MIIT), as well as in

- An April 2020 directive issued by the State Council and Central Committee of the Chinese Communist Party (CCP) designates data as the fifth factor of production—after land, labor, capital, and technology.[3] At the National People's Congress in March 2021, the outline of the 14th Five-Year plan called for "improving the market of data factors" (健全数据要素市场), and stressing the need to unlock the value of data to fuel the digital economy.[4]

- On November 30th of this year, China's Ministry of Industry and Information Technology released the 14th Five Year Plan (2021-2025) for China's big data industry. The plan defines big data as a strategic emerging industry, slated for greater state support to unlock the value of data. State supporting measures focus on expanding "international cooperation" between Chinese and foreign "big data services" companies in standard setting and research & development (R&D), and encourage multinationals to set up R&D centers in China. By 2025, the plan calls for China to set up new mechanisms to facilitate China's role in data trading and cross-border transfers. (建立数据资源产权、交易流通、跨境传输和安全等基础制度和标准规范) and "encourages Chinese firms to offer big data services in Belt and Road Initiative (BRI) countries and regions."

Beijing is also taking steps to centralize state control over data by breaking down silos or data islands across different government ministries and between the government and private companies, which have long plagued the government's ability to aggregate and coordinate data. Barriers to data sharing are due to a variety of reasons. Chinese companies are reluctant to share their data as valuable commercial intellectual property, while government agencies often push back against one another's access requests, guarding their data as a form of political power.[5]

An article by the Tencent Research Institute argues for facilitating more data flows to China's large tech platforms. Citing an International Data Corporation (IDC) estimate, the article states that "by

---

the Big Data Strategy (2017). The 13th Five Year Plan (2016–2020) calls for "fully implementing the promotion of the big data development initiatives and accelerating the sharing of data resources and development of applications, to assist in industrial transformation and upgrading . . . "

[3] Ouyang Shijia, "New guideline to better allocate production factors," April 10, 2020, China Daily, https://www.chinadaily.com.cn/a/202004/10/WS5e903fd7a3105d50a3d15620.html.

[4] Sina Online, "What Is the Meaning of the '14th Five-Year Plan' Outline (Draft) to Improve the Market of Data Elements? ("十四五"规划纲要（草案）提出健全数据要素市场有何深意)," March 5, 2021, https://finance.sina.com.cn/china/2021-03-05/doc-ikftssaq1688850.shtml.

[5] Yuan Yang and Nian Liu, "Alibaba and Tencent refuse to hand loans data to Beijing," Financial Times, September 18, 2019, https://www.ft.com/content/93451b98-da12-11e9-8f9b-77216ebe1f17; Martin Chorzempa, Paul Triolo, Samm Sacks, "China's Social Credit System: A Mark of Progress or a Threat to Privacy?" Peterson Institute for International Economics Policy Brief, June 2018, https://www.piie.com/publications/policy-briefs/chinas-social-credit-system-mark-progress-or-threat-privacy; Samm Sacks testimony before Senate Judiciary Committee hearing "Dangerous Partners: Big Tech and Beijing," March 4, 2020, https://www.judiciary.senate.gov/imo/media/doc/Sacks%20Testimony.pdf; Amba Kak and Samm Sacks, "Shifting Narratives and Emerging Trends in Global Data Governance Policy," AI Now and Yale Law School Paul Tsai China Center Policy Report, August 21, 2021, https://law.yale.edu/sites/default/files/area/center/china/document/shifting_narratives.pdf.

2025, the proportion of the world's data held by [China] will increase from 23.4 percent in 2018 to 27.8 percent, making China the first in the world. The open use of data resources will determine whether our country can seize the initiative in a new round of international competition and guarantee national data security through the development and growth of the digital industry."[6]

What are the implications for the United States of China's domestic and international efforts to acquire and make use of data as a strategic asset?

2. *Understanding China's motivations and different scenarios for how aggregated datasets could be used by the Chinese government is vital for creating effective U.S. policy.*

There are concerning potential uses of U.S. personal data from a national security perspective. Beijing is already presumed to have sensitive national security information from the theft of personnel records of roughly 21 million individuals from the U.S. Office of Personnel Management; travel information from a cyber attack on Marriott hotels covering roughly 400 million records; and credit data from Equifax on roughly 145 million people.[7] If additional sources of personal data such as location, social media, or pattern of life data were to be acquired or bought openly through unregulated data brokers and combined with what Beijing has already acquired through cyber theft, Chinese security services could use it to target individuals in sensitive government national security positions or military personnel for manipulation, blackmail, or other forms of coercion. This is particularly concerning from a counterintelligence perspective for individuals with security clearances or those with access to critical infrastructure.

As Chinese online services and network infrastructure gain in prominence around the world, it is also possible that the Chinese government could filter or monitor data processes abroad, just as the United States had done, as shown by Snowden, in utilizing data transmissions across U.S. networks for intelligence gathering. We also simply do not know what value and harm data created today will have in the future, regardless of who has access to it. As we move toward a world in which people have online profiles built on aggregated data, we must ask: what are the implications of the CCP gaining effective control of information flows beyond China's closed internet system? What are the implications as the CCP takes even more drastic steps to close off the loopholes that to this day keep even the Great Firewall relatively porous and circumventable[8] (e.g., stricter enforcement of

---

[6] Chen Weixuan et al., "Data Production Factors in the Framework of Macroeconomic Growth: History, Theory and Prospects (宏观经济增长框架中的数据生产要素：历史、理论与展望)," Tencent Research Institute, June 12, 2020, https://tisi.org/14625.

[7] "China's Collection of Genomic and Other Healthcare Data from America: Risks to Privacy and U.S. Economic and National Security," National Counterintelligence and Security Center Fact Sheet, February 2021.

[8] Magaret Roberts assesses China's Great Firewall relies on "friction-based censorship" that "works through distraction and diversion. It nudges — but does not force — most users away from unsavory material. This framing of censorship, Roberts says, helps explain why, even though China's Great Firewall is porous and can be circumvented, the number of people who 'jump the wall' using a virtual private network (VPN) remains relatively low. People are not necessarily afraid of legal or political consequences of using a VPN, but rather the process of doing so is deemed too bothersome or offers too little value for the effort in most people's day-to-day lives." Stanford Freeman Spogli Institute, March 6, 2020, https://fsi.stanford.edu/news/china%E2%80%99s-great-firewall-built-friction-based-censorship-says-margaret-roberts.

restrictions on virtual private networks [VPNs] or shifting from a blacklist to a whitelist approach to permissible websites so technical controls can keep pace with online content deemed threatening)?

At the same time, the Chinese government's use of data is not monolithic. Different actors are seeking data not just for security and surveillance, but also as fuel for the digital economy and other basic administrative functions. Outside observers of China often view Beijing's actions solely through the lens of security, neglecting the economic development drivers that play an important role. China's Data Security Law makes explicit that security and development must be balanced in China's data-governance system. These two competing priorities have shaped China's cyber bureaucracy for years. This long-standing internal source of friction and negotiation has contributed, at least in part, to the Chinese government not necessarily enforcing to date the strictest or most conservative security-oriented readings of Chinese cybersecurity laws and regulations. An entire early chapter of the Data Security Law was dedicated to this balance, indicating a recognition by Chinese authorities that state power hinges not only on security of data, but also on its commercial use, and that China must therefore find an effective way to leverage both at once. This duality also is driving an ambitious national effort to classify all data resources held by government and industry by category and grade ("categorized and graded protection system for data;" 数据分类分级保护制度). The goal is to distinguish less sensitive data for circulation to fuel the economy from data that should be locked down with tighter security restrictions.

As China grows in prosperity, and its leadership seeks to assert state control over data for both strategic and economic gain, the United States must also develop a comprehensive vision and regulation to maintain leadership. Leading Chinese data scholar Dr. Hong Yanqing writes that "China should also consider how to enable Chinese enterprises to control and use more data globally. After all, the United States can extend its 'arm' because its enterprises are all over the world." Hong observes that Chinese tech companies need access to global data flows, and that if the United States and the European Union are able to align on digital policies, China will be at a disadvantage of creating split products for different markets (for example, Bytedance segmenting its global and Chinese versions of the apps TikTok and Douyin). He adds that this approach "prevents Chinese ICT companies from upgrading services by using a global data pool and limits the gains from the economies of scale. Once the United States and the European Union reach an agreement, at least their enterprises can avoid data localization and segregating storage, which puts Chinese ICT enterprises at a disadvantage."[9]

Inaction by the United States will result in failure to create the interoperable coalition on data that Chinese leaders fear. Stalled progress on Privacy Shield and a global vision for data flows like APEC Cross-Border Privacy Rules underscore the challenges ahead.

[9] Hong Yanqing, "Game of Laws: Cross-Border Data Access for Law Enforcement Purposes," trans. Yale Law School Paul Tsai China Center, originally published in Global Law Review in Chinese. This article is the result of a special 2018 project by the Ministry of Justice, "Big Data and Cybersecurity Legislation" (18SFB1005), in which the author participated. https://law.yale.edu/sites/default/files/area/center/china/document/game_of_laws-7.pdf.

Recommendations

To be effective, U.S. policy should be based on an accurate understanding of why data matters. The analogy of data as the new oil is false, and leads to bad policy that treats data as a finite and zero-sum resource that is only valuable in large volumes. Matt Sheehan writes that five dimensions are crucial for machine learning data today: quantity, depth, quality, diversity, and access.[10] This understanding of data's value matters because it means that policies by Beijing or Washington that seek to hoard or wall off data as a national resource from the other could have unintended consequences that lessen national power, rather than increase it.

Lack of regulation in the United States makes Americans' sensitive data vulnerable to privacy and security harms not only from sophisticated state-backed cyber intrusions, but also from the unregulated industry of data brokers around the world trading in consumer data without transparency or controls. Setting basic standards on what data can be collected and retained by all companies will help protect U.S. personal data, regardless of where the risk originates. Developing a comprehensive federal privacy law that includes restrictions on data brokers is vital to this effort, along with the creation of strong enforcement mechanisms. Inaction by the United States means ceding leadership and influence in setting international standards to both Europe and China in setting international standards.

Without higher standards for data security and privacy, U.S. citizen data held by unregulated private companies are more vulnerable to breaches by hackers from China or from being sold to third-parties openly buying, aggregating, and selling consumer data. For example, Equifax's many security issues are well-documented, such as the company's failure to patch known vulnerabilities that ultimately left exposed the data of 145 million Americans. But the hack was also conducted by a foreign government entity with sophisticated hacking capabilities and access to considerable state resources. Companies should not have access to such a volume of personal data that it creates a target to be hacked or transmitted to China.

This reality is also why bans on Chinese software applications are not an effective way to secure Americans' data. Even if TikTok were American-owned, for example, it could still legally sell data to data brokers that could transmit it to China's security services.

Given this, American data is shockingly exposed and will remain that way so long as restrictions on data flows only focus on specific companies from countries deemed adversaries.

Debate over a range of issues will make progress on a federal privacy law slow. In the meantime, having baseline rules for the data broker industry would contribute to closing off vectors that make American's data vulnerable to exploitation by a range of actors.

---

[10] Matt Sheehan, "Much Ado About Data: How America and China Stack Up," Macro Polo, June 16, 2019, https://macropolo.org/ai-data-us-china/?rp=e.

We must also keep in mind that U.S. actions to respond to data security risks posed by the Chinese government are not occurring in a vacuum. Our policy approach should be tailored to take into account the fact that technology competition with China will not only play out in the United States and China, but also in other parts of the world from India to Europe. How we respond to Chinese companies operating in the United States has ramifications on whether other countries are willing to accept our vision of data governance.

The ability of U.S. firms to maintain a high rate of innovation depends upon access to global markets, talent, and, perhaps most important, datasets. But rising data sovereignty policies around the world are an increasing obstacle to the ability of U.S. companies to operate internationally, beyond China. These policies are an effort by nation-states to ensure control over data by prohibiting transfers of data out of the country or seeking to limit foreign access to certain kinds of data. In this context, U.S. actions will be a reference and a roadmap for other governments that are concerned about U.S. companies and the U.S. government getting access to their citizens' data.

The United States should work with like-minded governments to develop a common set of standards that would allow data to flow—building off of the concept of "data free flows with trust" put forward by Japan.[11] A multilateral approach should be based on creating a system of incentives for compliance. The United States could lead the way in setting up a certification system that would extend benefits to countries whose data regimes and companies meet certain clear criteria for data protection. The OECD privacy guidelines, for example, could serve as a reference in creating a baseline for commercial data flows.[12]

We need to address national security risks where they exist, but that should be done as one part of a broader U.S. initiative for comprehensive data privacy and higher cybersecurity standards for all companies —whether domestic or foreign. Failure to offer a compelling vision for U.S. data governance will make the United States less secure, less prosperous, and less powerful, and allow more space around the world for companies controlled by the CCP to flourish.

[11] Data Free Flow with Trust (DFFT): Paths towards Free and Trusted Data Flows, World Economic Forum, https://www.weforum.org/whitepapers/data-free-flow-with-trust-dfft-paths-towards-free-and-trusted-data-flows.
[12] "The OECD Privacy Framework," Organisation for Economic Co-operation and Development, 2013, https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.