Chinese Investment in the Data Economy

Capt Jesse Lubove

Air University Advance Research – ISR Considerations

31 March 2022

Abstract

China is one of the leaders of the internet data economy, and Chinese investments in this area have both commercial and national security implications for the U.S. With apps like TikTok and WeChat, China has collected vast amounts of internet data from millions of Americans. This data includes highly sensitive information such as biometric, demographic, health, location, and device data. The Chinese government, assisted by Chinese technology firms, is using this data to shape the information environment, improve its artificial intelligence (AI) technology, and increase its intelligence collection capabilities. There are almost no restrictions on the data Chinese companies can gather on Americans, but China has taken steps to protect the data of its citizens from American companies. To counter China's current asymmetric advantage in data, the U.S. government needs to strengthen U.S. data privacy laws and limit the export of internet data to China.

Chinese Investment in the Data Economy

China is one of the leaders in the internet data economy, and Chinese investments in this area have both commercial and national security implications for the U.S. With apps like TikTok and WeChat, China has collected vast amounts of internet data from millions of Americans. The Chinese government, assisted by Chinese technology firms, is using this data to shape the information environment, improve its artificial intelligence (AI) technology, and increase its intelligence collection capabilities. There are almost no restrictions on the data Chinese companies can gather on Americans, but China has taken steps to protect the data of its citizens from American companies. To counter China's current asymmetric advantage in data, the U.S. government needs to strengthen U.S. data privacy laws and limit the export of internet data to China.

Types of Chinese Investment in Data

TikTok, which is owned by Beijing-based ByteDance, has nearly 100 million active users in the U.S. TikTok collects detailed biometric, demographic, location, and device data from its users.[1] The WeChat messaging app, which is owned by Shenzhen-based Tencent, has an average of nearly 20 million active users in the U.S. Tencent can also gather extensive data from its users and read all messages sent through WeChat.[2] Alibaba, another Chinese tech giant, has a cloud computing service called Alibaba Cloud. Alibaba Cloud has over 4 million global customers including large American companies such as Ford, IBM, and Hewlett Packard.[3] Alibaba Cloud is currently under investigation by the U.S. Commerce Department for how it stores customer data

[1] "Terms of Service | TikTok," TikTok, February 2019, https://www.tiktok.com/legal/terms-of-service?lang=en.
[2] Fergus Ryan, Audrey Fritz, and Daria Impiombato, "TikTok and WeChat: Curating and Controlling Global Information Flows," TikTok and WeChat (Australian Strategic Policy Institute, 2020), 45, https://www.jstor.org/stable/resrep26120.
[3] Alexandra Alper, "U.S. Examining Alibaba's Cloud Unit for National Security Risks - Sources," *Reuters*, January 19, 2022, sec. Technology, https://www.reuters.com/technology/exclusive-us-examining-alibabas-cloud-unit-national-security-risks-sources-2022-01-18/.

and "whether the Chinese government could gain access to" private employee data and intellectual property.[4]

Chinese tech companies are not unique in collecting data on Americans. Most internet services are free because those services collect data from their users to sell targeted advertisements. However, data collection by Chinese companies is unique because of the close relationship between Chinese technology companies and the Chinese government. In 2015, China's Ministry of Public Security discussed plans for establishing government outposts at technology companies.[5] China's National Intelligence Law requires Chinese technology companies to censor politically sensitive content, assist with criminal investigations, and develop domestic surveillance technology.[6] In a similar manner, U.S. companies can be compelled by the U.S. to assist in criminal or national security investigations. However, the U.S. process is more transparent and has established legal processes for companies to challenge government requests in court.[7] In China, companies do not have similar rights to appeal Chinese government requests.

First-party data collection by Chinese-owned companies has come under increasing scrutiny in the U.S. In September 2020, the Trump administration attempted to ban TikTok and WeChat citing national security concerns.[8] However, there are other ways that Chinese companies gather data from Americans that have received less scrutiny. Chinese companies,

---

[4] Alper.
[5] Liza Lin and Josh Chin, "China's Tech Giants Have a Second Job: Helping Beijing Spy on Its People," *Wall Street Journal*, November 30, 2017, sec. Pro Cyber, https://www.wsj.com/articles/chinas-tech-giants-have-a-second-job-helping-the-government-see-everything-1512056284.
[6] Gregory C. Allen, "Understanding China's AI Strategy" (Center for a New American Security, February 6, 2019), https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.
[7] Ellen Nakashima, "Apple Vows to Resist FBI Demand to Crack IPhone Linked to San Bernardino Attacks," *Washington Post*, February 17, 2016, sec. National Security, https://www.washingtonpost.com/world/national-security/us-wants-apple-to-help-unlock-iphone-used-by-san-bernardino-shooter/2016/02/16/69b903ee-d4d9-11e5-9823-02b905009f99_story.html.
[8] Bobby Allyn and Bill Chappell, "U.S. To Bar Downloads Of TikTok, WeChat," *NPR*, September 18, 2020, sec. Technology, https://www.npr.org/2021/06/09/1004750274/biden-replaces-trump-bans-on-tiktok-wechat-with-order-to-scrutinize-apps.

such as Tencent, have bought internet data from third-party actors known as data brokers. Data brokers are companies that collect, bundle, and sell internet data to other organizations. Data from data brokers have many uses including marketing and intelligence. There are almost no U.S. government regulations for data brokers, and data brokers can be used to bypass other government restrictions. In 2019, the Committee on Foreign Investment in the United States (CFIUS) ordered Beijing Kunlun Tech to sell its majority ownership stake in Grindr, a popular gay dating app. CFIUS was concerned that the Chinese government could use the app to acquire sensitive data on Americans such as HIV status, dating habits, location history, and sexual orientation. In 2020, Beijing Kunlun Tech sold its stake in Grindr to a U.S.-based investment group.[9] However, the Norwegian government released a report that showed Grindr was still selling user data to 18 different third-party companies, including Tencent, and the sale of Grindr user data to Chinese companies could not be stopped by CFIUS.[10]

China has supplemented its legal investments in the data economy with illegal hacking. From 2014 to 2015, Chinese government actors hacked the U.S. Office of Personnel Management and stole the records of 4 million government employees including sensitive security clearance information.[11] In 2017, Chinese government actors stole the financial records of 145 million Americans from Equifax, a credit reporting agency.[12] In 2018, Chinese actors

[9] Yuan Yang and James Fontanella-Khan, "Grindr Sold by Chinese Owner after US National Security Concerns," *Financial Times*, March 7, 2020, https://www.ft.com/content/a32a740a-5fb3-11ea-8033-fa40a0d65a98.
[10] Justin Sherman and Kamran Kara-Pabani, "How a Norwegian Government Report Shows the Limits of CFIUS Data Reviews," *Lawfare* (blog), May 3, 2021, https://www.lawfareblog.com/how-norwegian-government-report-shows-limits-cfius-data-reviews.
[11] Ellen Nakashima, "Chinese Breach Data of 4 Million Federal Workers," *Washington Post*, June 4, 2015, sec. National Security, https://www.washingtonpost.com/world/national-security/chinese-hackers-breach-federal-governments-personnel-office/2015/06/04/889c0e52-0af7-11e5-95fd-d580f1c5d44e_story.html.
[12] Katie Benner, "U.S. Charges Chinese Military Officers in 2017 Equifax Hacking," *The New York Times*, February 10, 2020, sec. U.S., https://www.nytimes.com/2020/02/10/us/politics/equifax-hack-china.html.

stole guest data and passport information for 500 million people from the Marriott.[13] China has demonstrated its interest in acquiring data on Americans through any means possible. The next section will explore how China can use that data to harm U.S. national security interests.

Impacts of Data on U.S. National Security

Information Environment

Chinese companies have used their position in the data economy to shape the information environment to China's advantage. Researchers at the University of Toronto discovered that WeChat has an automatic filter that will block messages on controversial subjects, such as the 1989 Tiananmen Square massacre, even for users outside of China.[14] For example, Zhou Fengsuo, a Chinese-American pro-democracy activist, had his WeChat messages blocked and account suspended multiple times even though he is a U.S. citizen living in the U.S.[15] Chinese technology companies do not even have to remove a post or suspend a user to shape the information environment. In 2019, a VICE journalist posted multiple videos to TikTok that were critical of the Chinese government with the tag #Xinjiang. Tiktok allowed the videos to be posted, but the videos did not show up when users searched for #Xinjiang.[16] This more subtle version of censorship is harder to detect and can be more insidious since the user might not even recognize how their voice is being restricted.

Artificial Intelligence

The same companies that are major investors in the data economy are also leaders in AI technology. Investment in the data economy assists the development of AI because large, high-

[13] David E. Sanger et al., "Marriott Data Breach Is Traced to Chinese Hackers as U.S. Readies Crackdown on Beijing," *The New York Times*, December 11, 2018, sec. U.S., https://www.nytimes.com/2018/12/11/us/politics/trump-china-trade.html.
[14] Jason Q. Ng, "Tracking Censorship on WeChat's Public Accounts Platform" (University of Toronto, July 20, 2015), https://citizenlab.ca/2015/07/tracking-censorship-on-wechat-public-accounts-platform/.
[15] Ryan, Fritz, and Impiombato, "TikTok and WeChat: Curating and Controlling Global Information Flows," 25.
[16] Ryan, Fritz, and Impiombato, 15.

quality datasets allow developers to better train AI.[17] Additionally, better AI can help companies collect more data on their users which further increases these companies' advantages in AI. For instance, TikTok's success in the U.S. is largely attributed to its addictive AI recommendation algorithm, and because the app is so popular, TikTok continues to gather more user data that TikTok can use to increase the effectiveness of that algorithm.[18]

The same AI technology is often used for both commercial and security purposes in China. Alibaba uses its AI facial recognition to authorize mobile payments with its "smile to pay" program, and Alibaba also uses that same technology for China's smart surveillance cities.[19] In 2020, Alibaba Cloud even advertised how it could use its AI facial recognition software to distinguish between the faces of Uyghurs and Han Chinese to assist Chinese authorities in their repression of the Uyghurs in Xinjiang.[20]

Much of China's use of AI and data is focused on ensuring internal stability. However, the same information and technology can be used to identify U.S. intelligence personnel. China has already stolen biometric data from major international airports and transit hubs.[21] China can supplement hacked data with data from its technology companies. For example, TikTok changed its terms of service in 2021 to say that it could collect "faceprints and voiceprints" from its

---

[17] Kai Jia et al., "The Application of Artificial Intelligence at Chinese Digital Platform Giants: Baidu, Alibaba and Tencent," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, February 26, 2018), https://doi.org/10.2139/ssrn.3154038.
[18] Ben Smith, "How TikTok Reads Your Mind," *The New York Times*, December 6, 2021, sec. Business, https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html.
[19] Jia et al., "The Application of Artificial Intelligence at Chinese Digital Platform Giants"; Karen Gilchrist, "Alibaba Launches 'Smile to Pay' Facial Recognition System at KFC in China," *CNBC*, September 4, 2017, sec. Marketing.Media.Money, https://www.cnbc.com/2017/09/04/alibaba-launches-smile-to-pay-facial-recognition-system-at-kfc-china.html.
[20] Raymond Zhong, "As China Tracked Muslims, Alibaba Showed Customers How They Could, Too," *The New York Times*, December 16, 2020, sec. Technology, https://www.nytimes.com/2020/12/16/technology/alibaba-china-facial-recognition-uighurs.html.
[21] Zach Dorfman, "China Used Stolen Data to Expose CIA Operatives in Africa and Europe," *Foreign Policy*, December 21, 2020, https://foreignpolicy.com/2020/12/21/china-stolen-us-data-exposed-cia-operatives-spy-networks/.

users.[22] China can use this data with its facial recognition technology to target U.S. intelligence personnel around the world[23] In 2020, *Foreign Policy* reported that China used stolen data to identify undercover CIA officers in Africa and Europe.[24] This threat will continue to grow as China improves its AI technology and gathers more data.

Recommendations

To counter Chinese investment in the data economy, the U.S. needs to strengthen its privacy laws. Four U.S. states have passed their own laws with varying restrictions on the types of data companies may collect.[25] However, America has no comprehensive federal data privacy laws. There is a patchwork of federal legislation that covers specific sectors, like the Health Insurance Portability and Accountability Act (HIPAA), but almost all of the data currently collected by Chinese companies is not covered by any federal legislation.[26]

To address the security vulnerabilities mentioned in this paper, Congress should pass legislation to restrict the data that companies may gather on Americans and limit the sale of that data to foreign companies. Of the four state privacy bills passed, the California Consumer Privacy Act (CCPA) is the most comprehensive and should serve as a model for federal legislation. The CCPA requires transparency on the types of data companies collect on their users and allows Californians to limit the sale of their data to third-party data brokers.[27] If a

---

[22] Bobby Allyn, "Senators Demand TikTok Reveal How It Plans To Collect Voice And Face Data," *NPR*, August 18, 2021, sec. Technology, https://www.npr.org/2021/08/18/1028633650/senators-demand-tiktok-reveal-how-it-plans-to-collect-voice-and-face-data.

[23] Jenna McLaughlin and Zach Dorfman, "'Shattered': Inside the Secret Battle to Save America's Undercover Spies in the Digital Age," *Yahoo News*, December 30, 2019, https://news.yahoo.com/shattered-inside-the-secret-battle-to-save-americas-undercover-spies-in-the-digital-age-100029026.html.

[24] Zach Dorfman, "China Used Stolen Data to Expose CIA Operatives in Africa and Europe."

[25] Taylor Kay Lively, "US State Privacy Legislation Tracker," March 24, 2022, https://iapp.org/resources/article/us-state-privacy-legislation-tracker/.

[26] Thorin Klosowski, "The State of Consumer Data Privacy Laws in the US (And Why It Matters)," *Wirecutter: Reviews for the Real World* (blog), September 6, 2021, https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

[27] "California Consumer Privacy Act (CCPA)," State of California - Department of Justice - Office of the Attorney General, October 15, 2018, https://oag.ca.gov/privacy/ccpa.

similar framework was used at the federal level, the compliance burden on technology companies would be minimized because those companies already have to follow those limitations if they offer their services in California.

Additionally, Congress should put strict limits on the data that third-party companies may sell to foreign companies, especially to companies in China. The U.S. already limits the export of certain items based on national security grounds and could apply that approach to data. In April 2021, Senator Ron Wyden, a Democrat from Oregon, proposed the Protecting Americans' Data from Foreign Surveillance Act that would do that, but there has not been any movement on the bill since then.[28]

While the U.S. has failed to pass federal privacy legislation, China's Personal Information Protection Law (PIPL) went into effect on 1 November 2021. PIPL emphasizes the national security risks of data and attempts to limit China's exposure to those risks.[29] PIPL places restrictions on the data companies may collect and requires strict security measures for companies that move the data of Chinese citizens out of China. However, the law does not limit the Chinese government's access to data.[30] Unless action is taken, China will continue to benefit from unrestricted access to U.S. data while restricting the U.S.'s access to similar data on Chinese citizens. Moreover, China's PIPL could help shape the de facto global standard for data privacy while the U.S. is forced to play by other countries' rules. China is currently dictating the rules for the data economy, and without U.S. action, China's advantage in the information environment will continue to grow.

---

[28] Drew Harwell, "Wyden Urges Ban on Sale of Americans' Personal Data to 'Unfriendly' Foreign Governments," *Washington Post*, April 15, 2021, https://www.washingtonpost.com/technology/2021/04/15/personal-data-foreign-government-ban/.
[29] Matt Burgess, "Ignore China's New Data Privacy Law at Your Peril," *Wired*, November 5, 2021, https://www.wired.com/story/china-personal-data-law-pipl/.
[30] Eileen Yu, "China's Personal Data Protection Law Kicks in Today," *ZDNet*, October 31, 2021, https://www.zdnet.com/article/chinas-personal-data-protection-law-kicks-in-today/.

References

Allyn, Bobby. "Senators Demand TikTok Reveal How It Plans To Collect Voice And Face Data." *NPR*, August 18, 2021, sec. Technology. https://www.npr.org/2021/08/18/1028633650/senators-demand-tiktok-reveal-how-it-plans-to-collect-voice-and-face-data.

Alper, Alexandra. "U.S. Examining Alibaba's Cloud Unit for National Security Risks - Sources." *Reuters*, January 19, 2022, sec. Technology. https://www.reuters.com/technology/exclusive-us-examining-alibabas-cloud-unit-national-security-risks-sources-2022-01-18/.

Benner, Katie. "U.S. Charges Chinese Military Officers in 2017 Equifax Hacking." *The New York Times*, February 10, 2020, sec. U.S. https://www.nytimes.com/2020/02/10/us/politics/equifax-hack-china.html.

Bobby Allyn and Bill Chappell. "U.S. To Bar Downloads Of TikTok, WeChat." *NPR*, September 18, 2020, sec. Technology. https://www.npr.org/2021/06/09/1004750274/biden-replaces-trump-bans-on-tiktok-wechat-with-order-to-scrutinize-apps.

Burgess, Matt. "Ignore China's New Data Privacy Law at Your Peril." *Wired*, November 5, 2021. https://www.wired.com/story/china-personal-data-law-pipl/.

State of California - Department of Justice - Office of the Attorney General. "California Consumer Privacy Act (CCPA)," October 15, 2018. https://oag.ca.gov/privacy/ccpa.

Drew Harwell. "Wyden Urges Ban on Sale of Americans' Personal Data to 'Unfriendly' Foreign Governments." *Washington Post*, April 15, 2021. https://www.washingtonpost.com/technology/2021/04/15/personal-data-foreign-government-ban/.

Gilchrist, Karen. "Alibaba Launches 'Smile to Pay' Facial Recognition System at KFC in China." *CNBC*, September 4, 2017, sec. Marketing.Media.Money. https://www.cnbc.com/2017/09/04/alibaba-launches-smile-to-pay-facial-recognition-system-at-kfc-china.html.

Gregory C. Allen. "Understanding China's AI Strategy." Center for a New American Security, February 6, 2019. https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy.

Jenna McLaughlin and Zach Dorfman. "'Shattered': Inside the Secret Battle to Save America's Undercover Spies in the Digital Age." *Yahoo News*, December 30, 2019. https://news.yahoo.com/shattered-inside-the-secret-battle-to-save-americas-undercover-spies-in-the-digital-age-100029026.html.

Jia, Kai, Martin Kenney, Juri Mattila, and Timo Seppala. "The Application of Artificial Intelligence at Chinese Digital Platform Giants: Baidu, Alibaba and Tencent." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, February 26, 2018. https://doi.org/10.2139/ssrn.3154038.

Justin Sherman and Kamran Kara-Pabani. "How a Norwegian Government Report Shows the Limits of CFIUS Data Reviews." *Lawfare* (blog), May 3, 2021. https://www.lawfareblog.com/how-norwegian-government-report-shows-limits-cfius-data-reviews.

Liza Lin and Josh Chin. "China's Tech Giants Have a Second Job: Helping Beijing Spy on Its People." *Wall Street Journal*, November 30, 2017, sec. Pro Cyber.

https://www.wsj.com/articles/chinas-tech-giants-have-a-second-job-helping-the-government-see-everything-1512056284.

Nakashima, Ellen. "Apple Vows to Resist FBI Demand to Crack IPhone Linked to San Bernardino Attacks." *Washington Post*, February 17, 2016, sec. National Security. https://www.washingtonpost.com/world/national-security/us-wants-apple-to-help-unlock-iphone-used-by-san-bernardino-shooter/2016/02/16/69b903ee-d4d9-11e5-9823-02b905009f99_story.html.

———. "Chinese Breach Data of 4 Million Federal Workers." *Washington Post*, June 4, 2015, sec. National Security. https://www.washingtonpost.com/world/national-security/chinese-hackers-breach-federal-governments-personnel-office/2015/06/04/889c0e52-0af7-11e5-95fd-d580f1c5d44e_story.html.

Ng, Jason Q. "Tracking Censorship on WeChat's Public Accounts Platform." University of Toronto, July 20, 2015. https://citizenlab.ca/2015/07/tracking-censorship-on-wechat-public-accounts-platform/.

Ryan, Fergus, Audrey Fritz, and Daria Impiombato. "TikTok and WeChat: Curating and Controlling Global Information Flows." TikTok and WeChat. Australian Strategic Policy Institute, 2020. https://www.jstor.org/stable/resrep26120.

Sanger, David E., Nicole Perlroth, Glenn Thrush, and Alan Rappeport. "Marriott Data Breach Is Traced to Chinese Hackers as U.S. Readies Crackdown on Beijing." *The New York Times*, December 11, 2018, sec. U.S. https://www.nytimes.com/2018/12/11/us/politics/trump-china-trade.html.

Smith, Ben. "How TikTok Reads Your Mind." *The New York Times*, December 6, 2021, sec. Business. https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html.

Taylor Kay Lively. "US State Privacy Legislation Tracker," March 24, 2022. https://iapp.org/resources/article/us-state-privacy-legislation-tracker/.

TikTok. "Terms of Service | TikTok," February 2019. https://www.tiktok.com/legal/terms-of-service?lang=en.

Thorin Klosowski. "The State of Consumer Data Privacy Laws in the US (And Why It Matters)." *Wirecutter: Reviews for the Real World* (blog), September 6, 2021. https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

Xi Jinping. "'Strive to Become the World's Primary Center for Science and High Ground for Innovation.'" *DigiChina* (blog). Accessed March 29, 2022. https://digichina.stanford.edu/work/xi-jinping-strive-to-become-the-worlds-primary-center-for-science-and-high-ground-for-innovation/.

Yang, Yuan, and James Fontanella-Khan. "Grindr Sold by Chinese Owner after US National Security Concerns." *Financial Times*, March 7, 2020. https://www.ft.com/content/a32a740a-5fb3-11ea-8033-fa40a0d65a98.

Yu, Eileen. "China's Personal Data Protection Law Kicks in Today." *ZDNet*, October 31, 2021. https://www.zdnet.com/article/chinas-personal-data-protection-law-kicks-in-today/.

Zach Dorfman. "China Used Stolen Data to Expose CIA Operatives in Africa and Europe." *Foreign Policy*, December 21, 2020. https://foreignpolicy.com/2020/12/21/china-stolen-us-data-exposed-cia-operatives-spy-networks/.

Zhong, Raymond. "As China Tracked Muslims, Alibaba Showed Customers How They Could, Too." *The New York Times*, December 16, 2020, sec. Technology. https://www.nytimes.com/2020/12/16/technology/alibaba-china-facial-recognition-uighurs.html.