February 2021

CHINA'S COLLECTION OF GENOMIC AND OTHER HEALTHCARE DATA FROM AMERICA: RISKS TO PRIVACY AND U.S. ECONOMIC AND NATIONAL SECURITY

The National Counterintelligence and Security Center

Would you want your DNA or other healthcare data going to an authoritarian regime with a record of exploiting DNA for repression and surveillance? For years, the People's Republic of China (PRC) has collected large healthcare data sets from the U.S. and nations around the globe, through both legal and illegal means, for purposes only it can control. While no one begrudges a nation conducting research to improve medical treatments, the PRC's mass collection of DNA at home has helped it carry out human rights abuses against domestic minority groups and support state surveillance. The PRC's collection of healthcare data from America poses equally serious risks, not only to the privacy of Americans, but also to the economic and national security of the U.S.

The Value of Your DNA

- Your DNA is the most valuable thing you own. It holds the most intimate details of your past, present and potential future—whether you are prone to addiction or high-risk for cancer. It is your unique genetic code and can enable tailored healthcare delivery to you.

- Losing your DNA is not like losing a credit card. You can order a new credit card, but you cannot replace your DNA. The loss of your DNA not only affects you, but your relatives and, potentially, generations to come.

China Prioritizes the Collection of Healthcare Data

- The PRC views bulk personal data, including healthcare and genomic data, as a strategic commodity to be collected and used for its economic and national security priorities. (Genomic data is a broad term referring to your entire genetic sequence—all your DNA).

- The PRC is investing heavily in the "biotech revolution" and has enacted national policies prioritizing the collection of healthcare data both at home and abroad to achieve its goal of becoming a global biotech leader. It has designated biotech as a "strategic emerging industry" and prioritized state support for its biotech industry in national plans like the Made in China 2025 plan.[1]

- The PRC understands the collection and analysis of large genomic data sets from diverse populations helps foster new medical discoveries and cures that can have substantial commercial value and advance its Artificial Intelligence and precision medicine industries.

    - In 2016, the PRC announced a $9 billion, 15-year project to collect, analyze, and sequence genomic data to become a global leader in precision medicine—a process designed to provide tailored treatments based on the genetic makeup and environment and lifestyle of individual patients.[2]

- With the COVID-19 pandemic, the PRC aggressively marketed Chinese COVID-19 testing kits around the world, along with laboratories to support COVID-19 testing. By August 2020, China's leading genomics company, BGI, said it had sold test kits to 180 countries and established labs in 18 countries in the past six months.[3]

- According to the U.S.-China Economic and Security Review Commission, these COVID-19 labs have been providing Chinese researchers with access to healthcare data from around the globe.[4] Despite their aggressive pitches to U.S. states, there is no evidence Chinese companies have been able to establish such COVID-19 labs in the U.S.

China's Access to U.S. Healthcare Data

- Nevertheless, the PRC has for years been able to gain access to U.S. healthcare data, including genomic data, through a variety of channels, both legal and illegal.

- U.S. healthcare data may be particularly attractive and valuable to China because of the ethnic diversity of the U.S. population.[5] And compared to other nations, the U.S. has fewer safeguards on medical and healthcare data, including data for research purposes.[6] U.S. safeguards focus primarily on privacy, not national security, which creates a vulnerability for foreign actors to gain access to data on U.S. persons.

- Over the years, Chinese companies have taken advantage of this environment by investing in U.S. firms that handle sensitive healthcare and other types of personal data, providing them entry to the U.S. market and access to this data.

    - For instance, China's BGI purchased U.S. genomic sequencing firm Complete Genomics in 2013. In 2015, China's WuXi Pharma Tech acquired U.S. firm NextCODE Health to later form WuXi NextCODE Genomics.[7]

- Chinese companies have also gained access to U.S. healthcare data by partnering with hospitals, universities, and other research organizations in America. These U.S. entities routinely seek low-cost genomic sequencing services for their facilities, which Chinese biotech firms can often provide due to Chinese government subsidies. (In February 2020, BGI said it could sequence a human genome for just $100.[8]) These partnerships allow U.S. entities to expand their research capabilities, while Chinese firms gain access to more genetic data on more diverse sets of people, which they can use for new medical products and services.

    - A 2019 report found at least 15 Chinese companies that were either headquartered in China with a U.S. presence, or wholly located in China, were licensed to perform genetic testing or whole genomic sequencing on patients in the U.S. healthcare system, giving them direct access to the genetic data of patients in the U.S.[9]

- Finally, the PRC has gained access to U.S. healthcare data through illicit means, including theft of research and cyberattacks.

    - Among the most notorious examples was the 2015 hack of U.S.-based health insurer Anthem, Inc., in which data on some 78.8 million persons was stolen from Anthem's computer networks, including health identification numbers, names, Social Security numbers, employment and income data and other information. A U.S. Justice Department indictment in 2019 charged two individuals based in China for the hack of Anthem and three other U.S. companies.[10]

China's History of DNA Exploitation

- Concerns over the exploitation of healthcare and genomic data by the PRC are not hypothetical. The PRC has a documented history of exploiting DNA for genetic surveillance and societal control of minority populations in Xinjiang, China.[11]

- Specifically, the PRC government has established a high-tech surveillance system across Xinjiang, as part of a province-wide apparatus of oppression aimed primarily against traditionally Muslim minority groups. An initiative launched by the PRC government in 2014 has been used to justify the collection of biometric data from all Xinjiang residents ages 12 to 65. Authorities have collected DNA samples, fingerprints, iris scans, and blood types. The biometric data is linked to individuals' identification numbers and centralized in a searchable database used by PRC authorities.[12]

- Specific abuses by the PRC government as part of this effort include mass arbitrary detentions, severe physical and psychological abuse, forced labor, oppressive surveillance used arbitrarily or unlawfully, religious persecution, political indoctrination, and forced sterilization of members of minority groups in Xinjiang.[13] All told, the PRC government in Xinjiang has detained more than 1 million members of Muslim minority groups in internment camps for Communist Party indoctrination since 2017.[14]

- In July 2020, the U.S. Department of Commerce sanctioned two subsidiaries of China's BGI for their role in conducting genetic analysis used to further the PRC government's repression of Uyghurs and other Muslim minority groups in Xinjiang.[15]

    - Over the past decade, China's BGI has partnered with many research and healthcare entities in America to provide them with genomic sequencing services, while also gaining access to health records and genetic data on people in the U.S.[16]

Implications for Privacy and U.S. National Security

- China's access to U.S. healthcare and genomic data poses serious privacy and national security risks to the U.S.

    - Through its cyber intrusions in recent years, the PRC has already obtained the Personal Identifying Information (PII) of much of the U.S. population.

- Recent breaches attributed to the PRC government or to cyber actors based in China include the theft of personnel records of roughly 21 million individuals from the U.S. Office of Personnel Management; the theft from Marriott hotels of roughly 400 million records; the theft of data from Equifax on roughly 145 million people; and the theft of data from Anthem on roughly 78 million people.[17]

- Furthermore, under the PRC's national security laws, Chinese companies are compelled to share data they have collected with the PRC government. Article 7 of China's 2017 National Intelligence Law, for instance, mandates that all Chinese companies and citizens shall support, assist, and cooperate with Chinese national intelligence efforts, and guard the secrecy of any national intelligence work that they are aware of. There is no mechanism for Chinese companies to refuse their government's requests for data.

- The combination of stolen PII, personal health information, and large genomic data sets collected from abroad affords the PRC vast opportunities to precisely target individuals in foreign governments, private industries, or other sectors for potential surveillance, manipulation, or extortion.

    - For instance, vulnerabilities in specific individuals revealed by genomic data or health records could be used to help target these individuals.[18] Data associated with an embarrassing addiction or mental illness could be leveraged for blackmail. Combine this information with stolen credit data indicating bankruptcy or major debt and the tools for exerting leverage increase. Such data sets could help the PRC not only recruit individuals abroad, but also act against foreign dissidents.

Economic Implications for the United States

- Aside from these immediate privacy risks, China's access to U.S. health and genomic data poses long-term economic challenges for the United States.

- The PRC's acquisition of U.S. healthcare data is helping to fuel China's Artificial Intelligence and precision medicine industries, while the PRC severely restricts U.S. and other foreign access to such data from China, putting America's roughly $100 billion biotech industry at a disadvantage.

- Over time, this dynamic could allow China to outpace U.S. biotech firms with important new drugs and health treatments and potentially displace American firms as global biotech leaders.

- Although new medicines coming out of China could benefit U.S. patients, America could be left more dependent on Chinese innovation and drug development for its cures, leading to a transfer of wealth, co-opting of new businesses and greater job opportunities in China.[19]

Endnotes:

[1] Mark Kazmierczak and Thilo Haneman, "China's Biotechnology Development: The Role of U.S. and Other Foreign Engagement," Gryphon Scientific and Rhodium Group (prepared for the U.S.-China Economic and Security Review Commission), February 14, 2019, 36-38.

[2] Jennifer Schenker, "China Leaps Ahead in Precision Medicine," The Innovator News, August 27, 2019.

[3] Kirsty Needham, "Special Report: COVID Opens New Doors for China's Gene Giant," Reuters, August 5, 2020.

[4] 2020 Annual Report to Congress, U.S.-China Economic and Security Review Commission, 309.

[5] Mark Kazmierczak and Thilo Haneman, "China's Biotechnology Development: The Role of U.S. and Other Foreign Engagement," Gryphon Scientific and Rhodium Group (prepared for the U.S.-China Economic and Security Review Commission), February 14, 2019, 135.

[6] Ibid, 115.

[7] Ibid, 120.

[8] Antonio Regalado, "China's BGI says it can sequence a genome for just $100," MIT Technology Review, Feb. 26, 2020.

[9] Mark Kazmierczak and Thilo Haneman, "China's Biotechnology Development: The Role of U.S. and Other Foreign Engagement," Gryphon Scientific and Rhodium Group (prepared for the U.S.-China Economic and Security Review Commission), February 14, 2019, 124.

[10] U.S. Department of Justice press release, "Member of Sophisticated China-Based Hacking Group Indicted for Series of Computer Intrusions, Including 2015 Data Breach of Health Insurer Anthem Inc. Affecting Over 78 Million People," May 9, 2019.

[11] U.S. Department of State, U.S. Department of Treasury, U.S. Department of Commerce, U.S. Department of Homeland Security: Xinjiang Supply Chain Business Advisory, "Risks and Considerations for Businesses with Supply Chain Exposure to Entities Engaged in Forced Labor and other Human Rights Abuses in Xinjiang," July 1, 2020, 4.

[12] Ibid, 4.

[13] Ibid, 4.

[14] Ibid, 2.

[15] U.S. Department of Commerce press release, "Commerce Department Adds Eleven Chinese Entities Implicated in Human Rights Abuses in Human Rights Abuses in Xinxiang to the Entity List," July 20, 2020.

[16] Mark Kazmierczak and Thilo Haneman, "China's Biotechnology Development: The Role of U.S. and Other Foreign Engagement," Gryphon Scientific and Rhodium Group (prepared for the U.S.-China Economic and Security Review Commission), February 14, 2019, 122.

[17] U.S. Department of Justice press release, "Attorney General William P. Barr Announces Indictment of Four Members of China's Military for Hacking into Equifax," Remarks as Prepared for Delivery, February 10, 2020.

[18] Mark Kazmierczak and Thilo Haneman, "China's Biotechnology Development: The Role of U.S. and Other Foreign Engagement," Gryphon Scientific and Rhodium Group (prepared for the U.S.-China Economic and Security Review Commission), February 14, 2019, 133.

[19] 2020 Annual Report to Congress, U.S.-China Economic and Security Review Commission, 314.