FBI Counterintelligence Division

BOLT

Bureau Outreach & Liaison Tool

Use of China-Based Big Data and AI Capabilities Present Risks to US Data

17 March 2023

China-based technology companies offering Big Data and AI services may provide the People's Republic of China (PRC) the ability to collect sensitive corporate data in support of PRC national goals. Leveraging third party companies is a cost-effective means to utilize Big Data, however, under PRC laws, Chinese companies may be compelled to provide access to that data to PRC Intelligence Services upon request.

PRC Interest in Collecting Big Data. The PRC Government leverages Chinese corporations to achieve global dominance in Big Data and AI technologies.

- In 2013, Chinese President Xi Jinping declared, "the vast ocean of data, just like oil resources during industrialization, contains immense productive power and opportunities. Whoever controls big data technologies will control the resources for development and have the upper hand."
- The PRC's National Five-Year Plan (FYP) provides a whole-of-government roadmap to dominate big data in identified high priority technologies. The current 14th Five-Year Plan (2021-2025) strategic goals identifies big data, cloud computing, big Internet of things, industrial networks, blockchain, artificial intelligence (AI), virtual reality and augmented reality as key technology industries of the PRC's digital economy.

Resources:

- The Most Powerful Data Broker in the World Is Winning the War Against the US, New York Times, 30 November 2021
- "14th Five-Year" Plan for the Development of the Big Data Industry, Georgetown University, 10 February 2022

PRC National Security Laws compel cooperation with security and intelligence services.

If US corporations provide data to China-based technology companies to collect, store, or analyze proprietary information, PRC intelligence services can compel China-based technology companies to hand over corporate data, which can be used for intelligence and targeting purposes.

- Cyber Security Law of 2017 implicates companies must localize certain types of data held within China's borders, including the data of foreign companies working in undefined critical industries.
- National Intelligence Law of 2017 stipulates that citizens or private organizations must assist the Ministries of Public Security and State Security in carrying intelligence work.
- China's Personal Information Protection Law, passed in 2021, puts no restrictions on the PRC's ability to collect, access, or retain locally produced data. This law implicates centralize government control of data collected in China writ large, including that of foreign companies operating in China.

Resources:

- Who Benefits from China's Cybersecurity Laws?, Center for Strategic and International Studies, 25 June 2020
- The Real Danger of China's National Intelligence Law, 23 February 2019
- What You Need to Know About China's Intelligence Law that Takes Effect Today, Quartz, June 2017

PRC Use of Big Data. The PRC intelligence and security services purchase services from Big Data and AI technology companies ostensibly to conduct domestic and international surveillance and also to surveil its populace.

- The PRC maintains a countrywide network of government data surveillance services — called "public opinion analysis software" developed over the past decade and designed to collect data on domestic Internet users and media as well as foreign targets from sources such as Twitter, Facebook, and other Western social media, according to press reporting.
- According to Department of Justice indictments in 2019 and 2020, Chengdu 404, a local cybersecurity firm, constructed a "big data" repository tool known as Sonar-X that allowed users to search social media records that had been collected for individuals of interest, presumably for use by PRC intelligence. Sonar-X was used to find records related to individuals linked to various Hong Kong democracy and independence movements, a U.S. media outlet that reported on China's repression of Uyghurs, and a specific Tibetan Buddhist monk.

Resources:

- China harvests masses of data on Western targets, documents show, Washington Post, 31 December 2021
- 2022 Annual Report to Congress, U.S.-China Economic and Security Review Commission, November 2022
- Smart cities for an intelligent nation, China Daily Hong Kong Edition, 6 January 2020
- Solutions, PERCENT

FBI Reporting Interests

If you are able to answer any of the following questions related to suspicious PRC activities on big data and AI mergers and acquisitions or strategic partnerships, please contact your local FBI Private Sector Coordinator. The information you have may help the FBI better understand the challenges you face and address the threats more effectively.

- Does your company or organization utilize big data technology or AI application products or services from PRC-based companies? If so, which companies' products and/or services are you utilizing?
  - What are some of the non-proprietary terms included in the company's business contracts?
  - Where does the company assert that they store and/or process US-origin customer big data?
  - How did your company learn about the services offered by the company?
  - What claims did the company make in their big data and AI sales pitch to your company?
  - What information have you or your employees been required to provide to the Chinese company in order to comply with PRC laws
- Does your company have any presence in China? If so, has your business presence in China been the subject to onsite inspection by PRC law enforcement or intelligence and security services?
- What oversight and discretion over the conduct of the businesses in China or abroad does the PRC government intervene in or influence operations at any time?

FBI Counterintelligence Division

Bureau Outreach & Liaison Tool

BOLT Handling Instructions

Please refer to the below TLP guidelines for handling instructions. Any information in the open source references (hyperlinked) has no handling caveats, as it is publicly available.

Traffic Light Protocol (TLP) Definitions

Color	What it means
TLP:RED For the eyes and ears of individual recipients only, no further disclosure.	Recipients may not share TLP:RED information with anyone else. In the context of a meeting, for example, TLP:RED information is limited to those present at the meeting.
TLP:AMBER Limited disclosure, recipients can only spread this on a need-to-know basis within their organization and its clients. TLP:AMBER+STRICT Restricts sharing to the organization only.	Recipients may share TLP:AMBER information with members of their own organization and its clients, but only on a need-to-know basis to protect their organization and its clients and prevent further harm.
TLP:GREEN Limited disclosure, recipients can spread this within their community .	Recipients may share TLP:GREEN information with peers and partner organizations within their community, but not via publicly accessible channels. TLP:GREEN information may not be shared outside of the community.
TLP:CLEAR Disclosure is not limited.	Recipients can spread this to the world, there is no limit on disclosure. Subject to standard copyright rules, TLP:CLEAR information may be shared without restriction.

Community: Under TLP, a community is a group who share common goals, practices, and informal trust relationships. A community can be as broad as all cybersecurity practitioners in a country (or in a sector or region).

Organization: Under TLP, an organization is a group who share a common affiliation by formal membership and are bound by common policies set by the organization. An organization can be as broad as all members of an information sharing organization, but rarely broader.

Clients: Under TLP, clients are those people or entities that receive cybersecurity services from an organization. Clients are by default included in TLP:AMBER so that the recipients may share information further downstream in order for clients to take action to protect themselves. For teams with national responsibility this definition includes stakeholders and constituents.