



IARIW – CIGI 2023

# IARIW – CIGI 2023

Thursday, November 2 – Friday, November 3

## **Better Governance to Unleash the Value of Data: China's Practice of Building a Data Trading System**

Alex He

(Centre for International Governance Innovation)

Rebecca Arcesati

(Mercator Institute for China Studies)

Paper prepared for the Conference on The Valuation of Data November 2 – November 3, 2023

Session 8: Data Governance

Time: Friday, November 3, 2023 [1:15-2:15 EST]

## Better governance to unleash the value of data: China's practice of building a data trading system

By Alex He and Rebecca Arcesati

### Abstract

*China is on course to become the top data-producing nation in the world and the government has designated data as a new factor of production. Yet, we still have a limited understanding of the workings of one of the world's largest data markets. This paper explores how Chinese policymakers at the central and local levels are working to establish a more integrated and better regulated national data market, an important and understudied dimension of China's data governance and its approach to data as economic and innovation policy lever. We start by tracing the policy, regulatory and institutional context behind China's plans for a "data trading system 2.0," explaining how this fundamentally differs from similar efforts that were pursued in the past. We then delve into various experiments by China's new generation of data exchange platforms to tackle old problems, including difficulties in the determination of data ownership, data valuation, and trust-building between data providers and buyers. These efforts coexist with emerging experimentation and innovation regarding the trading of AI training data, cross-border data trading, and the marketization of public data. The empirical observations suggest that the state-centric features of China's data market may pose challenges to its further development.*

### Introduction

In July 2022, a hacker claimed they had obtained the personal information of 1 billion Chinese residents—name, address, birthplace, national ID number, mobile number, case details—from the official database of the Shanghai police, offering to sell it for 10 bitcoins. In a country where the state's draconian surveillance of its citizens has historically sidelined personal data protection, such leaks are commonplace.<sup>1</sup> China's data black market is huge, and one of the reasons in 2021 the Chinese government introduced major data laws, marking a major step forward in the construction of a data governance regime. Not only does an unregulated market threaten social stability—the top concern for an autocracy whose main preoccupation is staying in power—but it also undermines policymakers' plans to turbocharge the country's entire governance and economic structures through digital technologies.

A state-backed research center estimated total revenues of China's data market to rise from CNY 81.5 billion (USD 11.7 billion) in 2021 to CNY 198.9 billion in 2025.<sup>2</sup> China's government considers this data as an economic factor of production on par with land, labor, capital, and technology, and a foundation of national power and competitiveness. It wants companies domestic and multinational, state-owned and private, as well as government bureaucracies to share and trade more data with each other in a secure and orderly fashion. Through industrial policy, it is building a nationwide network of data centers, computing centers, and supercomputers to better allocate the resources and capabilities required to

---

<sup>1</sup> <https://www.merics.org/en/comment/shanghai-police-database-breach-exposes-lax-data-protection>

<sup>2</sup> CICS-CERT, White Paper on data Trading Platform 2022.

process data resources in strategic industries, such as artificial intelligence (AI).<sup>3</sup> One thing is clear: Beijing has a forward-looking strategy for data as a lever of economic and innovation policy.

Since 2014, dozens of online data trading platforms, or data exchanges, have mushroomed across China. These function as intermediary institutions where organizations can buy and sell data products like virtual assets, but also freely query some data sets and access related services such as cleaning, visualization, and desensitization. Products on these platforms run the gamut of training data for autonomous vehicles to corporate credit information. Up until recently, however, these pilots were empty shells, and few companies were interested in using them. Since 2021 China has tried again, with major cities like Beijing, Shanghai, and Shenzhen establishing data exchanges under stronger government guidance and within clearer legal boundaries. Meanwhile, new regulations, standards, and institutions are being established to oversee China's data resources and related transactions, including a dedicated national authority.<sup>4</sup>

With few exceptions, these experiments at marketizing data have been ignored by the academic literature as China's efforts to secure data have attracted more attention. This paper aims to fill this gap by exploring an important dimension of China's data governance, namely how policymakers are seeking to establish a unified national data market. We study the most institutionalized data marketplaces that have emerged in Beijing, Shanghai, Shenzhen, Guiyang, and Guangzhou, while also examining interesting practices in other active data exchanges across the country. Relying on publicly available information retrieved on the internet, we examined their business models, ownership structures, regulatory arrangements, product lists, and track record in brokering deals between data providers and buyers.

In our examination, we focus on how data exchanges are approaching pressing governance problems which are also familiar to other data markets around the world, and therefore have relevance beyond China. The first challenge is ownership of data, which has not yet been resolved by legislation and is particularly acute, considering the peculiar features of data as an economic resource. To overcome this, policymakers seem to be opting for a flexible system which guarantees the rights to process, use, and commercialize data products, facilitated by the deployment of new technologies, such as blockchain, which can solve issues of privacy and traceability. The second challenge is pricing data assets, which is being tackled through an experimental price formation mechanism jointly run by the central government and select data exchanges.

The third aspect, to which we devote the most space, is the trust problem of China's data market and a resulting mismatch between supply and demand. To solve this, data exchanges have considerably changed their business model from simple intermediaries to more sophisticated service providers and ecosystem orchestrators. In addition to actively brokering deals and supervising the whole transaction process, many exchanges have integrated third-party service providers for key functions, such as certifying buyers and sellers and ensuring that data products comply with security and personal information protection laws and regulations.

The paper then presents in-depth case studies which explore how China's data exchanges are becoming avenues for regulatory, technological, and institutional experimentation. The trading of AI training data, pilots that facilitate cross-border data transactions, and mechanisms for the controlled trading of data

---

<sup>3</sup> <https://www.merics.org/en/comment/oceans-data-lift-all-boats-chinas-data-centers-move-west>

<sup>4</sup> <https://www.reuters.com/world/china/china-form-national-data-bureau-2023-03-07/>

products based on public data represent emerging trends in China's data market and deserve further study.

Despite these promising developments, the limitations of China's state-driven data trading model are apparent. Most data exchanges are controlled by state-owned assets and the government plays a clear coordination role, either by introducing specific providers and buyers to the data exchanges or, in some cases, by directly brokering transactions. In other cases, data products being traded on the platforms originate from public-private partnerships. Although more and more private firms are joining in, a majority of data providers and buyers that are active on the surveyed data exchanges appear to be either state-owned enterprises or companies with strong government connections. Altogether, our findings suggest that the role of the market in China's model of data circulation is yet to be determined.

The paper proceeds as follows. We first present background information on China's digital economy and the recent regulatory overhaul of internet companies' data-related practices. The two subsequent chapters analyze the policy, regulatory and institutional context of China's push to build a more unified and efficient data market. We then turn to China's construction of a data trading system and examine how data exchanges are approaching issues related to data ownership, data pricing, and trust-building, while also discussing how their business models are changing in the process. We elaborate on emerging areas of institutional and regulatory innovation by presenting case studies of select data exchanges. We conclude by highlighting the significance of data marketplaces, also beyond China.

## **Background: China's burgeoning data economy**

China has a huge amount of data, so much so that one prominent AI investor once argued that it would confer a decisive advantage to the country in the competition with the United States to develop and deploy AI.<sup>5</sup> Frequent headlines stress China's internet population, the largest in the world with a penetration rate of 76 percent.<sup>6</sup> These netizens mainly surf the internet from their smartphone, buying grocery, making payments, chatting, ordering taxi rides, learning, consuming and producing entertainment, and advertising products through a vibrant ecology of apps and platforms. These activities generate vast troves of data about citizens and their behaviors, all the while the world's largest network of surveillance cameras watches their movements and 'smart' public services, from healthcare to the judiciary, collect and digitize more information. Per one estimate, in 2018 China produced 7.6 zettabytes of data and will account for 27.8 percent of the global total by 2025, surpassing the United States.<sup>7</sup> By 2025, Beijing wants to have the foundations of a "data-driven market system" to be in place, and the digital economy to grow from 7.8 percent to 10 percent of the nation's gross domestic product (GDP).<sup>8</sup>

But numbers are not everything—aspects such as quality, depth of coverage, and access matter. For example, China's ambitions to lead the world in AI require that relevant actors have access to suitable data. Data about people's train rides can be practically useless for training an algorithm to perform automated target tracking in military operations, for instance; conversely, that data may be crucial to a railway equipment manufacturer. China still faces challenges in getting good data where it is required. A

---

<sup>5</sup> Kai Fu-Lee. *AI Superpowers*

<sup>6</sup> Internet Development White Paper, 2023, <https://www.globaltimes.cn/page/202308/1297155.shtml>

<sup>7</sup> IDC and Seagate, <https://www.seagate.com/files/www-content/our-story/trends/files/data-age-china-idc.pdf>

<sup>8</sup> Digital Economy FYP

recent white paper by an influential government think tank detailed China's supply bottlenecks in AI training data, an increasingly pressing issue given the global emergence of GPT-4 and other large language models.<sup>9</sup> Additionally, more than 80 percent of China's information and data resources is said to be in the hands of government bureaucracies at all levels. Then Chinese Premier Li Keqiang cited the figure in 2016 when trying to push for public data opening to improve government services.<sup>10</sup>

At the same time, data sharing and circulation can raise thorny issues when personal and government data are involved. Having recognized the sensitivities and strategic significance of data, over the past three years Chinese policymakers engineered an unprecedented regulatory overhaul of the digital economy, aimed at ensuring its "healthy development".<sup>11</sup> Major new laws, supported by a growing body of implementing regulations and sectoral specifications, now govern how China's data is to be handled—the Data Security Law, Personal Information Protection Law, Cybersecurity Review Measures, and Outbound Data Transfer Security Assessment Measures. The motives are complex, ranging from citizens' outcry about tech firms' violations of their privacy to the Chinese Communist Party (CCP)'s concerns over national security and social stability.<sup>12</sup>

Underlying the regulatory frenzy is a sentiment that private firms' data monopolies are bad for China's overall socioeconomic development, and therefore the state must wrestle back control. A case in point, in July 2021 authorities launched a cybersecurity review into ride hailing giant Didi Chuxing following the company's rushed initial public offering (IPO) on the New York Stock Exchange. A year later, company was fined CNY 8 billion (USD 1.2 billion) due to its excessive and illegal collection of personal information,<sup>13</sup> though the decision was widely believed to reflect authorities' fear that US regulators could get their hands on data sensitive to China's national security.<sup>14</sup> Meanwhile, beginning 2021 Jack Ma's Ant Group was forced to restructure and share its consumer credit data with a joint venture partly owned by China's central bank.<sup>15</sup> Both Didi and Ant had reportedly resisted requests to share consumer data with the government.<sup>16</sup>

Firms also need access to government data. Indeed, the impact of such data access has been huge in the Chinese computer vision industry, where many facial recognition firms grew and innovate thanks to readily available biometric information about the populace which had been provided by public security organs.<sup>17</sup> The CCP's authoritarian governance has spurred the rise of a booming surveillance industry,

---

<sup>9</sup> CAICT, 2023 data factors white paper

<sup>10</sup> [https://china.chinadaily.com.cn/2016-05/14/content\\_25277793.htm](https://china.chinadaily.com.cn/2016-05/14/content_25277793.htm)

<sup>11</sup> <https://finance.sina.cn/2022-05-18/detail-imcwipik0418822.d.html>

<sup>12</sup> <https://www.technologyreview.com/2018/03/28/671113/chinas-citizens-do-care-about-their-data-privacy-actually/>; <https://carnegieendowment.org/2022/11/02/how-food-delivery-workers-shaped-chinese-algorithm-regulations-pub-88310>

<sup>13</sup> <https://digichina.stanford.edu/work/translation-chinese-authorities-announce-2b-fine-in-didi-case-describe-despicable-data-abuses/>; <https://merics.org/en/comment/didi-fine-marks-new-phase-beijings-rectification-tech-sector>

<sup>14</sup> <https://www.wsj.com/articles/in-the-new-china-didis-data-becomes-a-problem-11626606002>

<sup>15</sup> <https://www.wsj.com/articles/chinas-new-power-play-more-control-of-tech-companies-troves-of-data-11623470478>; <https://www.reuters.com/world/china/ant-groups-micro-loan-service-huabei-begins-share-data-with-chinas-central-bank-2021-09-22/>

<sup>16</sup> <https://www.ft.com/content/1651bc67-4112-4ce5-bf7a-d4ad7039e7c7>; <https://www.yicaglobal.com/news/didi-has-refused-to-provide-user-data-to-authorities-traffic-official-says>

<sup>17</sup> [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3679716](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3679716)

but in other sectors state organs and state-owned industries have been more jealous about their data. In addition to data access, tech firms also want the resulting data products they generate to be valued adequately. Efficient data marketplaces are thus expected to serve a range of functions for both business and government and to fuel China's data-driven economy.

## Literature review: Unleashing a new factor of production

In October 2019, the Fourth Plenum of the 19th CCP Central Committee designated data as a “factor of production” alongside land, labor, capital, and technology.<sup>18</sup> The definition, subsequently formalized by the CCP Central Committee and State Council in the April 2020 “Opinions on Building a More Perfect System for Market-based Allocation of Factors of Production,”<sup>19</sup> crystalized the evolving thinking in Beijing around the role of data as an economic input with immense potential to drive digital transformation, innovation, and growth. Already in December 2017, the General Secretary of the CCP and China's unelected President, Xi Jinping, referred to data as “a new production factor, a foundational and strategic resource and an important productive force”.<sup>20</sup>

Although this approach is familiar from economics literature, and other jurisdictions are working to harness the potential of public and privately held data for innovation and the benefit of society, the unique configuration of state-market relations in China sets its data strategy apart from those of other countries. In the Chinese politico-legal system, the CCP is above the law and the party controls and steers significant segments of economic production in a “socialist market economy”. The political economy of China's data market is therefore worth close examination.

Harnessing the value of data is a major theme in the Chinese government's digital strategy, linked to important public policy and economic objectives. Both academic and policy-oriented literature on Chinese state security and surveillance practices have illuminated the CCP's understanding of data as a strategic lever of political and social control—the foundation of China's turn towards “data-driven governance” under Xi Jinping.<sup>21</sup> From policing and politically-motivated surveillance to the social credit system and anti-Covid-19 measures, China's leaders believe that efficient governance and public services hinge upon deriving better insights from data.<sup>22</sup> In economic terms, Xi Jinping's vision for the digital transformation, encapsulated in the interlocking Big Data, Cyber Great Power, and Digital China

---

<sup>18</sup> 19<sup>th</sup> Plenum Communiqué

<sup>19</sup> 2020 Opinions

<sup>20</sup> [https://www.gov.cn/zhengce/2020-04/09/content\\_5500622.htm](https://www.gov.cn/zhengce/2020-04/09/content_5500622.htm);

<https://news.cctv.com/2017/12/10/ARTI3HNR1LMiMiNZKmr1NMD1171210.shtml>

<sup>21</sup> <https://www.aspi.org.au/report/engineering-global-consent-chinese-communist-partys-data-driven-power-expansion>; <https://merics.org/sites/default/files/2020-05/Programming%20China.pdf>; <https://merics.org/en/ccp-2021-smart-governance-cyber-sovereignty-and-tech-supremacy>; <https://www.aspi.org.au/report/mapping-chinas-tech-giants-supply-chains-and-global-data-collection-ecosystem>; <https://www.thechinastory.org/yearbooks/yearbook-2021-contradiction/chapter-1-smart-governance-smarter-surveillance/>

<sup>22</sup> <https://www.tandfonline.com/doi/abs/10.1080/10357718.2021.1956430>; <https://www.vodafone-institut.de/studies/data-driven-containment-of-covid-19-in-china/>;

<https://www.nytimes.com/2022/06/25/technology/china-surveillance-police.html>; Chin and Lin, *Surveillance State*; Brussee, *Social Credit*. Greitens has also examined the local politics of information integration as an instrument of social and political control in China:

[https://scholar.google.com/citations?view\\_op=view\\_citation&hl=en&user=xlhDVBOAAAAJ&citation\\_for\\_view=xlhDVBOAAAAJ:QIV2ME\\_5wuYC](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=xlhDVBOAAAAJ&citation_for_view=xlhDVBOAAAAJ:QIV2ME_5wuYC).

strategies,<sup>23</sup> foresees the increased penetration of data-driven technologies, such as big data analytics, the internet of things (IoT), cloud computing, fifth-generation network wireless technology (5G) and AI into the fabrics of the “real economy”—along with public administration, social life, cybersecurity and environmental management;<sup>24</sup> the goal is to upgrade agriculture and traditional industries to fuel innovation and modernize the country’s economic system.

To realize its potential, however, data must circulate and flow to those economic actors that have the need and the ability to generate value from it. Official pronouncements as well as authoritative commentaries by scholars affiliated with government-owned research institutions stress the need to “activate the factor value of data” by cultivating an efficient data trading market.<sup>25</sup> They believe that data has a multiplier effect on other factors of production, such as labor and capital, and that such potential is currently held back by insufficient marketization.<sup>26</sup> For this reason, “basic data systems” are regarded as the foundation of China’s digital development, on equal footing as infrastructure.<sup>27</sup>

This process of all-round digitization and digitalization, termed “informatization” in Chinese political discourse, represents one of the two pillars of China’s nascent data protection regime.<sup>28</sup> The other pillar, cybersecurity—encompassing the warding off of any threats or risks emanating from networks, hardware and software, all the way to content and data itself—is the key to understanding Beijing’s approach to the digital sphere because the CCP has linked it to its own survival.<sup>29</sup> This prioritization in Beijing’s cyber and digital policy perhaps explains why the literature on China’s data governance regime has placed a comparatively stronger focus on the security aspects. But as Xi Jinping once said: “Cybersecurity and informatization are a single body with two wings, the two wheels of a single drive,” and “without informatization there is no modernization”.<sup>30</sup> In fact, development and use of data is so important that China’s Data Security Law explicitly assigns equal importance to it as it does to data security.<sup>31</sup> Of course, what matters is whether policy implementation manages to reconcile these conflicting priorities or gets stuck in a deadlock.

Academic literature has only begun to examine China’s approach to data as a factor of production. In her examination of data politics in China, Liu takes the characteristics of data which have been explored in economics literature as a starting point to show how they affect China’s foreign economic relations. She argues that data’s non-rivalrous and partially excludable nature creates externality, commitment, and valuation problems.<sup>32</sup> From the perspective of valuation, Liu’s findings suggest that the value which China’s economy generates from data, such as welfare gains, cannot be captured by traditional metrics

---

<sup>23</sup> [https://www.gov.cn/zhengce/2021-12/01/content\\_5655197.htm](https://www.gov.cn/zhengce/2021-12/01/content_5655197.htm); [https://www.xinhuanet.com/politics/2014-02/27/c\\_119538788.htm](https://www.xinhuanet.com/politics/2014-02/27/c_119538788.htm); [https://www.gov.cn/zhengce/2023-02/27/content\\_5743484.htm](https://www.gov.cn/zhengce/2023-02/27/content_5743484.htm)

<sup>24</sup> <https://digichina.stanford.edu/work/translation-xi-jinpings-speech-to-the-politburo-study-session-on-the-digital-economy-oct-2021/>

<sup>25</sup> <http://theory.people.com.cn/n1/2021/0720/c40531-32162737.html>

<sup>26</sup> Ibid.

<sup>27</sup> Layout Digital China Construction

<sup>28</sup> Creemers, 2022; <https://digichina.stanford.edu/work/analyzing-chinas-2021-2025-informatization-plan-a-digichina-forum/>

<sup>29</sup> Gao, Henry. 2021. Data Regulation with Chinese Characteristics

<sup>30</sup> <https://www.newamerica.org/cybersecurity-initiative/blog/chinas-strategic-thinking-building-power-cyberspace/>

<sup>31</sup> DSL art. 12

<sup>32</sup> <https://link.springer.com/article/10.1007/s12116-021-09319-8>

like GDP. Chinese scholars too are paying close attention to China's efforts to promote its digital economy through data governance, and some are publishing their findings in English-language journals.<sup>33</sup> Although it hardly offers an impartial scholarly analysis, a volume by Daoli Huang, a researcher affiliated with China's Ministry of Public Security, presented the as-yet most comprehensive analysis in English of China's data trading market and compared it with relevant efforts in the United States and the European Union.<sup>34</sup> Huang finds that despite recent legal and regulatory developments, China's data trading remains highly fragmented and lacks a clear legislative basis, relying instead on regulatory exploration and "self-restraint" by local data exchanges. In particular, the question of data ownership is yet to be resolved.

Outside academic literature, in recent years several think tank reports and commentaries began examining and problematizing the Chinese government's approach to data as a factor of production.<sup>35</sup> Some scholars have highlighted the difficult balancing act between security and economic development imperatives within China's data governance regime, which may be said to roughly map onto the relationship between the party state and the market.

According to Lee, the DSL "entrenches a nation-wide bureaucratic focus on exploiting this resource by requiring the state to support measures for data-related standardization, research, education, innovation, infrastructure building and public service provision. These provisions support state interventions for market optimization..."<sup>36</sup> Sacks and Kak briefly explore the practice of Chinese data trading pilots within the broader context of data governance as an economic policy lever, noting that "state power hinges not only on security of data, but also on its commercial use".<sup>37</sup> One key function of these pilots, they argue, is to maximize state control over and visibility into national data resources, including those of private firms.<sup>38</sup> In studying China's debates around and policy framework for data governance, respectively, Boullenois and He both find that China is pursuing a state-centric approach at the expense of personal rights.<sup>39</sup> Boullenois also speculates about the possible configuration of data property rights in China, suggesting that "massive amounts of data may be routinely obtained from individuals and companies and be used without much restriction based on a broad understanding of the 'public interest'".<sup>40</sup>

This apparent contradiction—between data marketization and state control—is the premise of this paper. China's leaders believe that the "market-based allocation of factors" cannot take place safely and

---

<sup>33</sup> <https://www.tandfonline.com/doi/abs/10.1080/17538963.2022.2068833>;  
<https://www.sciencedirect.com/science/article/pii/S2773067023000067>;  
<https://dl.acm.org/doi/pdf/10.1145/3239542>; [https://link.springer.com/chapter/10.1007/978-3-031-44725-9\\_6](https://link.springer.com/chapter/10.1007/978-3-031-44725-9_6)

<sup>34</sup> Huang, Daoli. *Research on the Rule of Law of China's Cybersecurity*.

<sup>35</sup> [https://law.yale.edu/sites/default/files/area/center/china/document/shifting\\_narratives.pdf](https://law.yale.edu/sites/default/files/area/center/china/document/shifting_narratives.pdf);  
<https://lillianli.substack.com/p/abridged-data-as-a-factor-of-production>;  
<https://digichina.stanford.edu/work/china-wants-to-put-data-to-work-as-an-economic-resource-but-how/>;  
<https://merics.org/en/comment/china-activates-data-national-interest>;  
<https://www.iss.europa.eu/content/chinas-data-strategy>

<sup>36</sup> <https://www.ifri.org/en/publications/notes-de-lifri/asie-visions/cyberspace-governance-china-evolution-features-and-future>

<sup>37</sup> [https://law.yale.edu/sites/default/files/area/center/china/document/shifting\\_narratives.pdf](https://law.yale.edu/sites/default/files/area/center/china/document/shifting_narratives.pdf)

<sup>38</sup> Ibid.

<sup>39</sup> <https://www.cigionline.org/publications/state-centric-data-governance-in-china/>

<sup>40</sup> Boullenois, Camille. 2021. China's Data Strategy.



efficiently if left unsupervised. In the socialist market economy, the state sees itself as the chief broker of data transactions. As Li has observed, designating data as a factor of production implies a belief that “data’s value to a nation is underpriced and currently subjected to market distortions.” The special attributes of data as a resource make it a form of semi-public good whose benefits may not accrue to society without government intervention.<sup>41</sup> The party state has long viewed the hoarding of data by government bureaucracies, state-owned enterprises (SOEs) and private firms as causing data silos, or “data islands”—effectively, market failures hampering economic development and efficient governance. This explains recent moves to break the data monopolies of internet companies, as well as longstanding attempts at making SOEs and government agencies open their data resources.

China’s bid to orchestrate a state-led data market emanates from a strategy that treats data as a national asset and an important economic input, emphasizing the role of the government in overseeing and administering its collection, processing, and organic distribution to all parts of public administration, society and the economy that need it. Matching supply and demand through an efficient data trading market is a key piece of the puzzle. If it succeeds, China’s experiment will have far-reaching implications not only for its economic and technological future, but also for foreign actors’ interaction with its burgeoning digital economy.

This paper fills a gap in the literature by advancing an empirical review of the structure and practices of data trading platforms in China, highlighting both the challenges they are facing in terms of data valuation and other issues in the economics of data, as well as the approaches and solutions they have put forth. Based on a detailed analysis of the websites of active data exchanges in China, combined with a close examination of government policies, regulatory documents, and authoritative reports and commentaries by Chinese experts, academics, and policymakers, this paper offers the first systematic analysis of China’s government-sanctioned data trading market as it undergoes unprecedented transformation, regulatory adjustment, and institutionalization.

## **Institutionalization and the making of a state-led data market**

This paper argues that China’s government is working on unifying and regulating the national data market in such a way as to have state-backed institutions control and manage data transactions. In other words, the state intervenes to unlock the value of data as an economic factor of production, while ensuring the protection of national security and personal information. The ongoing institutionalization of the data trading market is an important manifestation of this effort. Although local pilots for data marketplaces have existed across China for almost a decade, private ownership in some cases, coupled with a generalized lack of supervision, did not sit well with the government’s more recent resolve to crack down on monopolistic behaviors, leaks, theft, and misuse of citizens’ data by private actors.<sup>42</sup>

The DSL and PIPL, which both came in force in the second half of 2021, signaled regulators’ intention to put an end to this chaotic situation.<sup>43</sup> The PIPL regulates the relationship between data handlers and

---

<sup>41</sup> <https://lillianli.substack.com/p/abridged-data-as-a-factor-of-production>

<sup>42</sup> <https://tech.huangqiu.com/article/43ICh0ssQax>; <https://www.scmp.com/tech/big-tech/article/3127485/build-digital-china-country-must-first-deal-its-rampant-black-market>; <https://www.reuters.com/world/asia-pacific/china-ministry-orders-38-apps-rectify-excessive-collection-personal-data-2021-11-03/>; <https://technode.com/2019/03/18/china-surveillance-data-security/>

<sup>43</sup> DSL and PIPL

consumers, for example by obliging platforms to allow users to opt out of AI-enabled microtargeting and personalized recommendation. The DSL establishes a comprehensive architecture of interlocking systems aimed at securing all of China’s data, personal and non-personal, from theft, illegal appropriation, manipulation, and misuse. At the heart of this edifice is the hierarchical classification of data into different categories, according to its importance for national security. Sectoral and local regulators are responsible for this process, meaning that authorities are to decide what kinds of data can circulate in the market and on which conditions. Importantly, the DSL also codifies the official blessing of a “data exchange market”: In principle, data that is not personally identifiable or deemed sensitive for national security, the public interest, or China’s socioeconomic development may flow and be monetized within China’s market and across borders.<sup>44</sup>

China’s data market has since entered a new phase, with major cities including Beijing, Shanghai, Shenzhen, Guiyang, and Guangzhou establishing data exchanges under strong central guidance and tight supervision. The new marketplaces are meant to facilitate the “orderly sharing of data” to increase productivity and public welfare while protecting security and privacy.<sup>45</sup> From a fragmented and unregulated market, data trading in China is undergoing a profound transformation towards a better governed and more integrated construct. Given the government’s anxiety around lack of supervision, not only are the new exchanges required to comply with relevant legal and regulatory requirements, but they are also more tightly controlled by the state through both ownership arrangements as well as institutional backing.

Multiple policy documents have been released to guide this transformation, including opinions issued by the CCP Central Committee and State Council in April 2022 which stressed the importance of basic systems, such as rights protection, to creating a unified technology and data market.<sup>46</sup> The involvement of multiple government agencies in promoting this, in line with their respective and sometimes overlapping responsibilities, is visible in policymaking.<sup>47</sup> The 14th Five-Year Plan for National Informatization (14th Informatization FYP)—the master plan for China’s digital transformation until 2025—attaches great importance to building basic systems, governance structures, institutional mechanisms, and standards for optimizing the circulation and allocation of “data factors”.<sup>48</sup> The document calls for research into and definition of data property rights, as well as systems for valuation, pricing, transaction tracing, security inspection, and dispute resolution among others. Besides promoting the exploitation of public data resources, enterprises are encouraged to share their datasets. Still, until recently details on how all this is supposed to work in practice were conspicuously absent.

---

<sup>44</sup> Art. 19 of the DSL further states that “the state is to establish and complete management systems for data transactions, regulating data exchange conduct, and cultivating data exchange markets”. Art. 33 stipulates that “in the provision of services by institutions engaged in data transaction, intermediary services shall require the party providing the data to explain its origin and shall review and verify the identities of both parties to the transaction, storing records of the verifications and transactions”.

<sup>45</sup>

<https://ex.chinadaily.com.cn/exchange/partners/82/rss/channel/cn/columns/sn19a7/stories/WS5f446834a31030182d641111.html>

<sup>46</sup> <https://archive.ph/ly1ba>.

<sup>47</sup> 14<sup>th</sup> FYP for the Development of Big Data, 14<sup>th</sup> FYP for Digital Economy Development, and Digital China Construction Plan.

<sup>48</sup> <https://digichina.stanford.edu/work/translation-14th-five-year-plan-for-national-informatization-dec-2021/>

Then, in December 2022, the Chinese Communist Party’s Central Committee and the State Council released major opinions on creating an efficient and compliant system of data circulation and trading.<sup>49</sup> Among other things, the opinions (dubbed Data Twenty Measures in Chinese media and expert discussions) call for building “a system of data property rights with Chinese characteristics,” starting with the creation of a graded system through which rights pertaining to three distinct categories of data—public data, corporate data, and personal data—are to be established, verified, and enforced. More specifically, data rights shall be divided into three categories, namely the rights to possess data resources, to process and use data, and to commercialize related products. Moreover, policymakers seek to encourage experimentation (e.g., around pricing models) and promote a fair and efficient distribution of revenues among all market participants, also to stimulate large tech firms to share and trade their prized datasets.

Even before the central government took the initiative, some local governments had started formulating relevant policies and regulations. The first local law seeking to regulate data trading was promulgated in 2016 in Guizhou.<sup>50</sup> The municipalities of Shanghai and Shenzhen pioneered China’s first local data regulations, which both entered into force in January 2022.<sup>51</sup> Guangdong government policies introduced a number of innovations for managing data transactions in the province, such as the position of “government chief data officer” to coordinate the utilization of public data across government departments and digital spaces where licensed enterprises may train personal data in a secure manner.<sup>52</sup> Several exchanges have also introduced their own rules, covering issues ranging from catalogues of data prohibited from trading to standards for transactions.<sup>53</sup> A number of national standards were also released over the past years to flesh out more specific technical and regulatory requirements for the operations of data exchanges.<sup>54</sup>

The creation in March 2023 of a new bureaucratic agency to oversee and manage national data resources goes in the same direction. An institutional outcome of the 2023 National People’s Congress, the National Data Bureau (NDB) will sit under the authority of the National Development and Reform Commission (NDRC), China’s central macroeconomic planning agency under the State Council. While the CAC remains firmly in charge of security and data protection (in collaboration with other competent organs) the NDB will be responsible for tasks such as planning and coordinating the digital transformation of public services, society and the economy as well as managing China’s data resources to promote their use and circulation.<sup>55</sup> The rationale behind its creation is twofold. First, it clarifies the division of labor between data policymaking and regulation for economic growth on the one hand, and data security work on the other. Previously, the NDRC, CAC, Ministry of Industry and Information

---

<sup>49</sup> Data Twenty Measures; <https://merics.org/en/merics-briefs/data-market-foreign-talent-bio-based-materials>

<sup>50</sup> Regulations on Promoting the Development and Application of Big Data in Guizhou Province

<sup>51</sup> Shanghai data regs, Shenzhen data regs

<sup>52</sup> [https://www.gov.cn/xinwen/2022-09/23/content\\_5711256.htm](https://www.gov.cn/xinwen/2022-09/23/content_5711256.htm)

<sup>53</sup> Shanghai, Anhui, Guiyang, Harbin, Central China Big Data Exchange, Shenzhen, Tianjin (add sources)

<sup>54</sup> E.g. GB/T 37728-2019; GB/T 36343-2018; 信息安全技术数据交易服务安全要求

<sup>55</sup> <https://npcobserver.com/wp-content/uploads/2023/03/%E5%9B%BD%E5%8A%A1%E9%99%A2%E5%85%B3%E4%BA%8E%E6%8F%90%E8%AF%B7%E5%AE%A1%E8%AE%AE%E5%9B%BD%E5%8A%A1%E9%99%A2%E6%9C%BA%E6%9E%84%E6%94%B9%E9%9D%A9%E6%96%B9%E6%A1%88%E7%9A%84%E8%AE%AE%E6%A1%88.pdf>; the original Chinese name is 国家数据局, which may also be translated as “administration”: <https://digichina.stanford.edu/work/translation-establishing-the-national-data-administration-march-2023/>

Technology (MIIT), and other ministries shared coordinating responsibilities for the data economy, which caused chaos and inefficiency. Second, the NDB addresses the regional fragmentation that has long hampered China's efforts to make use of its rich data resources, especially those that sit with government departments and state-owned enterprises. Up until the reform, responsibility was divided among about 15 local administrations with little to no coordination.<sup>56</sup>

Tighter supervision has also hit cross-border data flows. In line with its "Dual Circulation" strategy for economic development, which seeks to rebalance China's economy away from global integration and towards greater self-reliance,<sup>57</sup> the Chinese government is prioritizing the domestic circulation of data – as well as the algorithms, patents, information and other digital goods it underpins – while profiting from cross-border flows in a very controlled manner.<sup>58</sup> As mentioned, Chinese leaders have diagnosed a complex and hostile international environment, where adversaries—first and foremost the United States—can exploit China's external connections to undermine its security. As a result, sovereign state control over China-origin data is viewed as a necessary measure for improving the resilience of, and building a security shield around, its digital as well as real economy. The logic is "local storage, outbound assessment".<sup>59</sup>

Strict localization requirements have been in place since the entering into force of the Cybersecurity Law in 2017, while some sectoral restrictions to data exports existing in areas like healthcare and finance were already in place.<sup>60</sup> But following recent developments, beside other mechanisms for outbound personal data transfer that were introduced in the PIPL and in a number of implementing regulations data traders must familiarize themselves with a strict and still opaque security review process led by the CAC. Such a review is triggered whenever a firm seeks to provide abroad either large quantities of personal data, or any amounts of so-called "important" data, a category which is yet to be clearly defined despite years of lobbying by domestic and foreign companies.<sup>61</sup> Adding complexity, in December 2022 the MIIT released rules for trial implementation, requiring industrial and telecom data processors to receive an authorization from the MIIT before exporting important or "core" data.<sup>62</sup> This complexity is by design: The DSL directs each industry regulator to draw up catalogues of important and core data within their respective jurisdictions, in order to define which data can be openly shared and exported.

However, there are signs that Chinese policymakers recognize the burden of these data export restrictions and are even considering a relaxation. Notwithstanding the prioritization of national security in its digital strategy, China's government has realized that ensuring efficient data circulation and unlocking the value of data will be vital for the country's future economic growth and technological innovation. The Data Twenty Measures are the latest and most detailed top-level policy guidelines to

---

<sup>56</sup> <https://www.jnexpert.com/article/detail?id=4260>

<sup>57</sup> "This model entails engaging international capital, financial, and technological markets when advantages can be gained while simultaneously bolstering indigenous capabilities to avoid overreliance on the global economy—due to national security concerns or the vagaries of global economic cycles". <https://www.csis.org/analysis/dual-circulation-and-chinas-new-hedged-integration-strategy>

<sup>58</sup> <https://www.21jingji.com/article/20220412/herald/21013a2c7571cab199f3fede25385303.html>

<sup>59</sup> Liu J. China's data localization

<sup>60</sup> Creemers, 2022

<sup>61</sup> The most recent guidelines on this matter is still in draft form:

<https://mp.weixin.qq.com/s/MqmdsxcCASX9Sh2gAZIGJw>

<sup>62</sup> MIIT rules,

[https://wap.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2022/art\\_e0f06662e37140808d43d7735e9d9fd3.html](https://wap.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2022/art_e0f06662e37140808d43d7735e9d9fd3.html)

that extent. In fact, at least in theory, Chinese leaders have long believed that security and development should mutually reinforce each other. Given that recent and ongoing institutional reforms have led to the emergence of a new wave of data exchanges and other types of data trading platforms in China, it is worth looking back and examining the progress that has been made to date, as well as the problems that are still lingering.

## **China's practice of building a data circulation and trading system: A mix of progress and problems**

The first wave of data trading platforms in China emerged in 2015, when Xi Jinping's administration released its national strategy for big data development.<sup>63</sup> It is telling that planning documents released between 2015 and 2016 already mentioned that China should pilot data trading platforms and set up systems for data asset registration, pricing, and property rights protection, among others.<sup>64</sup> However, dozens of data trading platforms, trading centers, and data exchanges established since 2014-2015 have mostly failed to meet expectations. As of 2020, China's first big data exchange, the Guiyang Big Data Exchange of Guizhou province—sometimes dubbed 'the big data valley of China'—had an annual trading volume of less than CNY 5 million, far less than the target of CNY 20 billion.<sup>65</sup> According to official figures, in 2021 data exchanges only accounted for 2 percent of China's overall data trading activity.<sup>66</sup>

During this period, China's data exchanges were, essentially, empty shells, with barely any data being traded. Numerous problems have marred these platforms, such as difficulties in pricing data products and defining the associated ownership and usage rights, a lack of appropriate approaches and standards for data valuation, the absence of mechanisms to facilitate trust-building among trading partners, missing links with third-party providers for key services like security audits and dispute arbitration, and until recently the absence of laws and regulations in the areas of data security and personal information protection. Unsurprisingly, the data black market has thrived: Researchers at Peking University's National School of Development estimated that it would reach a scale of CNY 150 billion in 2021.<sup>67</sup>

In a speech he gave in 2022, Wang Jiandong, an official from the Big Data Development Department of the NDRC-affiliated State Information Center, made clear that regional and bureaucratic fragmentation remains a major challenge. Authorities hope to overcome this with new plans for a "data trading system 2.0".<sup>68</sup> Data exchanges are expected to comply with China's young data laws and to operate under stronger government guidance. Consequently, more than ten new data trading platforms, centers and exchanges have been established since 2021.<sup>69</sup> (See Appendix for more details). Developments so far point to a mixed picture of progress and setbacks.

---

<sup>63</sup> Big data strategy included in government work report in 2014, then in 2015 China released the "Action Plan on Promoting Big Data Development".

<sup>64</sup> Big data strategy included in government work report in 2014; 2015 Action Plan on Promoting Big Data Development; 13th FYP National Informatization.

<sup>65</sup> [https://stock.stcn.com/djld/202107/t20210712\\_3426536.html](https://stock.stcn.com/djld/202107/t20210712_3426536.html)

<sup>66</sup> [https://www.sohu.com/a/681146049\\_121269250#google\\_vignette](https://www.sohu.com/a/681146049_121269250#google_vignette)

<sup>67</sup> <https://36kr.com/p/1947394695301768>

<sup>68</sup> [https://www.ndrc.gov.cn/wsdwhfz/202209/t20220913\\_1335479.html?code=&state=123](https://www.ndrc.gov.cn/wsdwhfz/202209/t20220913_1335479.html?code=&state=123)

<sup>69</sup> CAICT. 2023. 数据要素白皮书(2022 年) [White paper on data factor 2022], January.

<https://dl.ofweek.com/2023-07/ART-2022112-8420-30601885.html>

For starters, the data trading volume has increased at a decent pace. The major data exchanges in Shenzhen, Guiyang, and Guangzhou have reached a trading volume of more than CNY 1 billion yuan each as of mid-2023. The increased activity probably has to do with the newly established exchanges being in economically advanced areas of China, coupled with the increased institutionalization of the country's data economy. Some are piloting new rules and practices in data pricing, ownership, and trading, including around cross-border data transfers in the case of Shenzhen and Shanghai.

Yet despite this progress, the activity of these intermediaries is still limited compared to the huge size of China's big data industry and total data trading, most of which is carried out through direct over-the-counter trading between data providers and buyers. One plausible explanation is that some of the old problems that held back the development of China's data trading market in the past are still lingering. In fact, government policies like the Data Twenty Measures encourage exchanges to experiment with various solutions, for example in the areas of data ownership, pricing mechanisms, trading rules, standards for data valuation, as well as around the business models of these data marketplaces.

One big challenge that some data exchanges have developed regulatory and technological solutions for addressing is the determination and confirmation of data ownership rights. The main idea is to temporarily put aside the issue data ownership and instead guarantee the rights to benefit from a given dataset through use, processing, and commercialization. This approach already appears to have stimulated an increase in data trading volume. New technologies such as privacy-enhancing technologies (PETs), federated learning,<sup>70</sup> and blockchain are greatly helping promote data trading because they allow for a separation of ownership from processing and usage rights.

Another innovation has been the creation of a data price formation system through a cooperation between the government and data exchanges. The NDRC's Price Monitoring Center issued guidelines for data pricing and is working with data exchanges to establish data pricing mechanisms. While this debate is far from resolved, Chinese policymakers and subject-matter experts are leaning towards a cost pricing approach for data resources such as datasets and APIs, where inputs such as labor, time and equipment constitute the main costs associated with collecting and developing data resources. The situation differs for data assets, including most data products and services: Like stocks in finance, these assets embody some costs, but their value typically is highly customized and depends on the income expectation of data buyers and providers. In this case, income pricing is preferred because it accounts for factors such as historical prices, supply and demand, customer segmentation, and the level of model contribution.<sup>71</sup>

A more fundamental challenge is a lack of incentives for companies to put their data in the market for trading, which translates into a supply bottleneck. To remedy this, many data exchanges are trying to foster trustworthy ecosystems comprising data providers, buyers, as well as third-party institutions that provide services such as compliance, auditing, certification, insurance, asset evaluation, dispute settlement, training, and infrastructure support. The idea is to establish an environment conducive to mutual trust between buyers and sellers.

---

<sup>70</sup> A decentralized machine learning approach to train algorithms via multiple, independent datasets.

<sup>71</sup> The economic value of data needs to be evaluated in specific use cases. In a modeled use case, the economic value of data can be measured based on the level of contribution of data to the increased accuracy of a model. See <https://juejin.cn/post/7249361518831370299>.

Partly because of ecosystems becoming more diverse, the business models of data exchanges are also changing. From intermediary platforms that make money by collecting a commission fee, China's data marketplaces are evolving into more sophisticated entities that provide a wide range of data value-added services, such as data processing, data valuation, anonymization, compliance consulting, security certification, and technical support. With the release of the Data Twenty Measures, particularly a requirement for separation between data providers and exchanges, the latter are transferring most of these data value-added services to professional third-party providers. In this way, the exchanges themselves only handle basic services, such as registering market entities, data products and data services for trading, supervising regulatory compliance, and facilitating transactions.

Chinese data exchanges are not only tackling old problems, though. Some have become hotbeds for institutional, technological, governance, and regulatory experimentation and innovation. This is probably the most interesting aspect of China's efforts to create a thriving data market. Our research has identified three emerging trends.

First, catalyzed by the enthusiasm around generative AI products such as ChatGPT, strong demand for training data for AI models is stimulating growth in China's data trading market. Data providers such as DataTang, platform companies such as Baidu, Tencent and Netease, and AI companies like iFlytek all have their crowdsourcing platforms for data collection and annotation (labelling)—key steps in preparing data for AI training. Data exchanges in Beijing and Shanxi have begun carrying out AI training products trading.

Second, government policies encourage the market-based circulation of verified products, services and algorithmic models based on public data, where ensuring high standards of personal information protection and data security is particularly important. Authorities have been promoting open government data on different levels of governments since 2015, but most of these projects have stalled due to poor data quality, infrequent updates, and lack of standardization. The development of public data is expected to lead the growth of China's data trading market. The Beijing International Data Exchange has ventured into this area by trading data products based on public data.

Third, as discussed in the previous chapter, China's localization policies and regulatory requirements signify rather complicated procedures for domestic and foreign actors that wish to transfer data out of the country. The DSL's emphasis on the "secure and free flow of data across borders" implies that permissible flows are what remains once sweeping national security-motivated controls have been applied.<sup>72</sup> This is particularly the case for personal information and so-called 'important data', where authorities are combining new mechanisms for outbound data transfers, such as CAC-mandated security reviews, with the piloting of safe interfaces, or ports, through which cross-border data trading can take place safely and in compliance with China's law and regulations. The Shenzhen Data Exchange is a frontrunner in this area.

Underlying all these developments is the question of which relationship between market and state best serves the needs of China's new phase of data circulation. Most data exchanges either controlled by the government via a state-owned-assets holding, or 100 percent owned by state assets. A few of them exist through mixed ownership arrangements between the state and private companies. The average registered capital is between CNY 50 million and CNY 100 million, with the lowest at CNY 0.3 million and

---

<sup>72</sup> See Article 11 of Data Security Law of China.

the highest at CNY 800 million (specifically the Shanghai Data Exchange).<sup>73</sup>As subsequent chapters will demonstrate, Beijing's attempt at creating a state-led data economy is not without contradictions.

Before delving into emerging issues and trends in China's data valuation and circulation and reflecting on future trajectories, the following chapter will examine how, with the development of the second wave of China's data marketplaces since 2021-22, the major and more institutionally advanced data exchanges of Guiyang, Shenzhen, Shanghai and Beijing have been testing new solutions to the problems that characterized the country's data trading trials in the early days. Specifically, key challenges include the determination of data ownership rights, data pricing, as well as the shifting role of data exchanges from mere intermediaries to platforms seeking to attract more providers and buyers to the market by orchestrating trustworthy ecosystems.

## **Key issues in China's construction of a data trading system**

### **Data ownership**

China's 14th Five-Year Plan listed among its policy goals the creation of a legal regime for data property rights.<sup>74</sup> Ownership is highly difficult to determine due to digitalization and data's properties of a non-rivalrous and partial excludable resource—a semi-public good—with increasing returns to scale as well as certain externalities, like privacy, and an ambiguous allocation of property rights among consumers and firms.<sup>75</sup> For example, data controlled by digital platforms usually comprises data generated by platforms themselves, individual-level data such as consumers's behavioural data, and data generated by other businesses. Since the same data can accrue benefits to, as well as possessed, used, and disposed by more than one entity at the same time, the definition of property ownership defined in China's Civic Code, which implies absoluteness and exclusivity, does not apply.

As a result of the uncertainty surrounding data ownership, until recently data transactions in China were left in a legal limbo. The 2022 Data Twenty Measures marked an important step forward by dividing the legal rights of participants in data production, collection, processing, circulation and use into three categories: ownership of data resources, rights to process and use, and rights to commercialize data. The exact definition and scope of these three categories are not yet clear, nor is the feasibility of the whole construct in a politico-legal system that prioritizes state power over the rights of individuals and businesses.<sup>76</sup> What matters in the context of this analysis is that 2022 measures did open the door to the creation of data ownership rules and systems on the basis of which data can be legally traded in China.

Following this impetus, the data exchanges in Guiyang, Shanghai and Shenzhen all introduced data ownership registration systems to determine different rights associated with the data being traded on their platform, in addition to market entity registration for providers and buyers. The Shanghai Data

---

<sup>73</sup> White Paper on Data Trading Platform 2022

<sup>74</sup> FYP; Informatization FYP

<sup>75</sup> <https://www.cpb.nl/sites/default/files/omnidownload/CPB-Background-Document-Policy-Options-Data-Economy-Literature-Review.pdf>

<sup>76</sup> For further reference on this subject, see Alex He (2023) who argues China's state-centric data governance regime prioritizes a dual goal of bolstering both economic growth and national security at the expense of personal information protection, Henry Gao (2021) who emphasizes that the key to understand data regulation in China is 'security', and Sourabh Gupta (2023) believes China data regulators are trying to balance control and security with privacy, inclusion and commerce.



Regulations stipulate the legal property rights acquired for using and processing data.<sup>77</sup> The Shenzhen Data Exchange issued an Interim Measure for the Administration of Data Ownership Registration in July 2023, which appears to be the first normative legal document on data ownership registration in China. The measure contains detailed and comprehensive regulations on data registration management and is expected to have a ripple effect in other provinces and cities. The Guiyang Big Data Exchange started issuing different certificates for the three categories of rights outlined above: ownership, processing and use, and commercialization. Following a similar approach, the Shandong Exchange Platform established a registration system to clarify the source and use cases of data products, thus paving the way for a determination of data ownership.<sup>78</sup>

The certificates issued by the data exchanges can be used as a legal basis for data trading, as well as for other purposes such as financing and debt repayment, incorporating data assets into balance sheets, accounting, and dispute resolution.<sup>79</sup> This pragmatic approach could incentivize more companies to buy or sell data via institutional exchanges, for example by guaranteeing the protection of the property rights and interests of data processors such as digital platforms who invest sheer resources in collecting and processing data.<sup>80</sup> With the help of blockchain technology, each certificate becomes a traceable way of identifying the source and right holders associated with any given data product being traded on the exchanges. Leaving aside a discussion of the important privacy and consumer rights implications in China's politico-legal context, which falls outside the scope of this paper, the thorny issue of data ownership determination is set aside for the moment while the rights to process, commercialize and use data are guaranteed.

However, enforcing personal data protection throughout the whole data trading process is still a problem. This is mainly being tackled through the application of privacy-enhancing computing technology (PET), federated learning, and blockchain technology. The goal is threefold. First, the objective is to achieve a situation where "data being traded can be used but not seen,"<sup>81</sup> while also supervising the data trading process by tracing the source and transfer history of the traded data and controlling its final use. Second, digital technologies can technically enable a separation between data ownership rights, rights to process and use, and rights to commercialize. Third, data owners and processors can be granted different levels of control, such that the latter can only access the information required for processing and using the data. Experimentation with these technologies is underway on the exchanges of Beijing, Guiyang, Shenzhen, and Shanghai.

### Data pricing

In a future where data will be capitalized and securitized for investment, the pricing model China may take is worth studying. Missing unified pricing standards and mechanisms, the task of putting a price tag

---

<sup>77</sup> Regulations of Shanghai Municipality on Data

<sup>78</sup> White Paper on Data Trading Platform 2022

<sup>79</sup> [http://www.sz.gov.cn/cn/xxgk/zfxxgj/zcfg/content/post\\_10692613.html](http://www.sz.gov.cn/cn/xxgk/zfxxgj/zcfg/content/post_10692613.html)

<sup>80</sup> [www.allbrightlaw.com/SH/CN/10475/b0be235dd460442.aspx](http://www.allbrightlaw.com/SH/CN/10475/b0be235dd460442.aspx)

<sup>81</sup> Du, Chuan. 2022. "Data trading 2.0 coming and Privacy-Enhancing Computing makes data can be used but not seen [数据交易 2.0 时代来临, 隐私计算让数据“可用不可见”], Yicai [第一财经], June 7, <https://m.yicai.com/news/101436079.html>; Zhang, Ye. 2023 "Privacy-Enhancing Computing makes data can be used but not seen [隐私计算: 让数据“可用不可见”], *Science and Technology Daily* [科技日报], April 10, <http://finance.people.com.cn/n1/2023/0410/c1004-32660651.html>

on data has largely been left up for negotiation between providers and buyers. Combined with the scenario-based, highly customized features of data transactions, this easily leads to chaos and extortion in data pricing. Large digital platforms, for example, tend to charge a higher price based on their vast data power.<sup>82</sup> As discussed earlier, data monopolies have been a major target of regulators' rectification campaign in China's digital economy.

Both government officials and representatives of professional associations have come up with recommendations for data valuation and pricing. Wang Jiandong, who assumed the Deputy Director of the NDRC's Price Monitoring Center in 2023, advocated using cost pricing for data resources and an income-based approach for data assets. Specifically, the cost pricing approach considers all types of investment such as labor, time, and equipment in data collection and standardization, plus data quality and privacy. Based on all these inputs, the cost pricing approach can basically reflect the value of data resources. For data assets, an income pricing-oriented mechanism should be created to capture the expected income from future value.<sup>83</sup> Furthermore, the China Appraisal Society issued guidance on data asset evaluation in 2019, which did not differentiate data resources from data assets and introduced three approaches to pricing: cost pricing, income pricing, and market pricing. They suggest appraising the costs, expected future income, and historic prices as a basis for data assets pricing.<sup>84</sup>

Data exchanges in Guiyang, Shenzhen, and Shanghai have introduced third-party recommendations or guidelines in this regard. After data providers have made an initial offer, the data exchanges or third-party agencies set a reference price, considering both the costs embedded in the data as well as the possible benefits that buyers could derive from using it, plus other factors such as consumers' expectations, supply and demand, historic prices, and customer segments. This raises the more fundamental question of whether data exchanges should intervene in pricing. Officials from some of the major data exchanges seem to agree that some intervention is appropriate as it helps increase the data transaction rate.<sup>85</sup> However, a full consensus has not been reached. Whereas most products and services on the Guiyang Big Data Exchange have a clearly marked price, those on the Shenzhen Data Exchange are marked as "negotiable".<sup>86</sup>

Importantly, the central government seems to favor the exploration of data pricing formation mechanisms through close cooperation between policy research and the practices of data exchanges. The Pricing Monitor Center, whose main duty is to monitor prices across China and creating a price index for important goods and services, began working with exchanges in 2023, signing strategic cooperation agreements with exchanges in Shanghai and Guiyang. In Shanghai, this led to the establishment of a dedicated joint lab. In Guiyang, the Center directed the establishment of the first calculation program for data trading. With the Fujian Big Data Exchange, the center initiated a service center for data assets evaluation. The center also got involved when the Shenzhen Data Exchange promoted the issuance of a data asset-based credit line by a bank.

### **Business models and supervisory roles of data exchanges**

---

<sup>82</sup> [www.allbrightlaw.com/CN/10475/7b4fe58ed4a1e455.aspx](http://www.allbrightlaw.com/CN/10475/7b4fe58ed4a1e455.aspx)

<sup>83</sup> <https://news.cnstock.com/industry,rdjj-202304-5052479.htm>

<sup>84</sup> <http://www.cas.org.cn/docs/2020-01/20200109165641186518.pdf>

<sup>85</sup> <https://www.ccf.org.cn/YOCSEF/Branches/Shenzhen/News/lt/2023-05-25/791854.shtml>

<sup>86</sup> See the websites of Guiyang Big Data Exchange and Shenzhen Data Exchange.

A crucial difference between China's data exchanges 2.0 and those that emerged during the first wave lies in their ownership and business model. The new data exchanges are either state holdings or controlled 100 percent by state assets. The Guiyang Big Data Exchange, the first-ever data exchange in China established in 2015, also underwent restructuring from private control to 100 percent state-asset controlled after its disappointing performance. Thanks to such a strong government backing and rapid commercialization of new technologies such as PET, federated learning, and blockchain, the new exchanges have evolved into full-fledged service providers helping reduce barriers to data trading, including issues related to compliance, security, efficiency, and trust.

An examination of the websites and reports of the major data exchanges in Shenzhen, Guangzhou, Shanghai, and Guiyang revealed that the services provided typically encompass provider and buyer accreditation and certification, data registration, compliance (security, personal information protection, etc.), technical support, matchmaking between supply and demand, and even rule making, standard setting, and cross-border trading in some cases. This is in stark contrast with the first generation of data marketplaces in China, which merely acted as intermediaries between providers and buyers—with scarce success.

Compliance and security risk assessment are the most important and basic services since they are instrumental for creating a trustworthy trading environment. Some data exchanges combine compliance and security risk assessment, while others keep them separate. Security risk assessments focus on national laws and regulations—chiefly the DSL and PIPL—as well as applicable provincial and local government regulations, other regulations issued by industrial and sectoral regulators, and technical standards. Data exchanges set rules and procedures for compliance and security risk assessment, which they either carry out themselves or contract to third parties, usually law firms.

A typical example of the data trading process illustrates the rich variety of services offered by data exchanges:

- i) Setting requirements, standards, and procedures for data products and services, such requirements for data masking, encryption, anonymization, processing, and cleansing;
- ii) Conducting compliance, security risk and other kinds of assessment, typically through professional third-party service providers, following which data products and services may be certified and approved for trading;
- iii) Setting a reference price for data products and services and, upon confirmation by the providers, listing them on the exchange;
- iv) Displaying data products and services along with information such as a basic overview, description of the use cases, and modalities for bidding;
- v) Reviewing buyers's qualifications, monitoring the data transactions, and confirming the transactions.

While the basic process is the same, business modes can take many forms. In general, however, commission fees have proved ineffective, and most exchanges have abandoned them. Instead, income from the data value-added services provided and membership fees are becoming the main sources of revenue for most exchanges. Still, profitability remains a considerable challenge. With access to abundant capital thanks to government involvement, the newly established data exchanges can afford

not to be profitable for some time. The Shanghai Data Exchange and Zhengzhou Data Exchange Center even position themselves as quasi-public services institutions, not seeking profitability as the goal and only charging for the services provided to maintain their operation.<sup>87</sup> The long-term viability of this model, however, remains to be seen.

In any case, supervising trading itself is another non-trivial function of data exchanges, as it is a precondition to a regulated and efficient market. Indeed, the absence of proper regulations and standards is widely seen in China as the main reason behind the chaotic and inefficient development of data marketplaces over the past decade. The establishment of regulated exchanges with strict oversight over the whole trading process puts every link of the chain under proper supervision. Of course, new digital technologies greatly assist the exchanges in guaranteeing robust privacy and security standards, essentially enabling the creation of control rooms through which the whole process of data trading can be monitored. The idea is that all data traded on the exchanges can be traced in terms of its sources, past transfers, controller, and buyer, and any violations can be held accountable.

In this regard, it is significant that the Shenzhen Data Exchange, Guiyang Big Data Exchange, and Shanghai Data Exchange all issued their own rules for data trading and circulation.<sup>88</sup> Shenzhen also introduced comprehensive regulations to govern the participation of data providers and third-party service organizations in the data trading market.<sup>89</sup> These rules and regulations are supposed to cover all aspects of the data trading process, fulfilling the responsibilities of data exchanges as official, institutionalized supervisors of data transactions.

Clear rules are only part of the picture, however: Companies and public institutions need incentives to put their valuable datasets in the market for trading. As mentioned in earlier sections, only a fraction of China's data trading takes place via data exchanges.<sup>90</sup> Off-the-counter transactions dominate China's data trading market and are expected to continue to do so, which leads to a chaotic situation due to the lack of legal and security supervision. Trading via data exchanges with government endorsement is supposed to solve the problems and risks facing off-the-counter transactions. However, the lack of trust in the data exchanges among data providers and buyers is hindering their development.

Data in greatest demand sits in diverse sectors like healthcare, transportation, electricity, aviation, and digital platforms. Data providers also include specialized data analytics companies such as ShujuTang, whose main business revolves around either collecting, cleansing and processing data to build new products or taking orders directly from clients to provide customized solutions. On the demand side are large commercial banks, government agencies, AI companies and other high-tech enterprises with high demand for data, such as pharmaceutical companies. The following chapter will delve into how Chinese data exchanges are striving to match supply and demand through building trust among data buyers and providers.

---

<sup>87</sup> See the introduction at Shanghai Data Exchange at: <https://www.zzbdex.com/about>

<sup>88</sup> See the regulations for data trading in Shenzhen at:

[http://www.sz.gov.cn/cn/xgk/zfxxgj/zcfg/content/post\\_10454883.html](http://www.sz.gov.cn/cn/xgk/zfxxgj/zcfg/content/post_10454883.html); See the measures for data trading places in Shanghai at: <https://app.sheitc.sh.gov.cn/sjxwxgwj/694679.htm>; See the regulations for data trading and circulation in Guizhou at:

[https://dsj.guizhou.gov.cn/zwgk/gzhgfwxwjsjk/gfwxwjsjk/202212/t20221226\\_77732169.html](https://dsj.guizhou.gov.cn/zwgk/gzhgfwxwjsjk/gfwxwjsjk/202212/t20221226_77732169.html)

<sup>89</sup> <https://finance.sina.com.cn/tech/roll/2023-03-06/doc-imyixums0539941.shtml>

<sup>90</sup> see footnote 66.

## How China's data exchanges try to build trust

Practitioners involved in China's main data exchanges regard lack of trust as one of the key reasons why data providers and buyers chose to trade and transfer data outside the formal trading channels provided by data exchanges, therefore this issue deserves deeper discussion. To overcome this, several exchanges have started to prioritize the creation of trustworthy ecosystems for data transactions. Providers in sectors where a lot of data is produced, from utilities to internet platforms, as well as major buyers such as commercial banks, government agencies and AI companies, are incorporated into the ecosystem alongside third-party service providers who perform functions such as data quality certification, security inspection, and dispute resolution.

Meanwhile, the concept of "data brokers" has been practiced in Guangdong province since 2022 to facilitate trust-building in the data market.<sup>91</sup> These are defined as companies instead of natural persons in the data market, similarly to the securities companies in the securities market. However, the concept is quite opaque and associated with various and distinct actors, with the first batch of data brokers designated in Guangdong even including data providers who were registered on data exchanges. More data brokers are encouraged to be established to constitute one important link in the ecosystem of trust in data trading market.<sup>92</sup>

Creating such ecosystems of trusts has been a clear focus of China's major data exchanges. The Guiyang Big Data Exchange, for example, tried to train and introduce data providers from diverse fields: government, finance, healthcare, tourism, labor employment, telecommunications, electricity, transport, and meteorology. The Shanghai Data Exchange introduced the concept of "data dealers", which includes data brokers and a variety of data trading services providers. In Beijing, the local data exchange took a use case-based approach to set priorities for data trading for each exchange. This resulted in the prioritization on AI training data trading and circulation, since most large AI model companies in China are in Beijing and the municipal government also introduced relevant policies to encourage the development of AI training data.<sup>93</sup>

What follows are a few examples of such trust-building efforts by China's major data exchanges.

### Guaranteeing compliance and data quality

Shenzhen Weiyang Tech, an AI infrastructure provider that registered as a data provider on the Shenzhen Data Exchange, received an unsecured credit enhancement loan worth CNY 10 million from China Everbright Bank's Shenzhen branch based on the company's data products listed on the exchange.<sup>94</sup> The ecosystem built by the Shenzhen Data Exchange, which includes third-party service providers that assist

---

<sup>91</sup> [www.21jingji.com/article/20220810/herald/16cf0bfa0bdec64a428abc28273494c0.html](http://www.21jingji.com/article/20220810/herald/16cf0bfa0bdec64a428abc28273494c0.html)

<sup>92</sup> [https://www.sohu.com/a/695817609\\_398084](https://www.sohu.com/a/695817609_398084)

<sup>93</sup> [www.cbdio.com/BigData/2023-06/12/content\\_6173931.htm](http://www.cbdio.com/BigData/2023-06/12/content_6173931.htm);  
[https://www.beijing.gov.cn/fuwu/lqfw/ztzl/gdec2023/qyjs/202307/t20230704\\_3154181.html](https://www.beijing.gov.cn/fuwu/lqfw/ztzl/gdec2023/qyjs/202307/t20230704_3154181.html)

<sup>94</sup> Tang, Wei. 2023. "The first data asset-based unsecured credit enhancement loan in the country issued in Shenzhen [全国首笔! 无质押数据资产增信贷款深圳落地]", April 4, *证券时报网(STCN)*, [www.stcn.com/article/detail/833337.html](http://www.stcn.com/article/detail/833337.html); Zhu, Lin. 2023, "China Everbright Bank's Shenzhen branch works together with Shenzhen Data Exchange to successfully finish the first data asset financing for small and micro businesses. [光大银行深圳分行携手深圳数据交易所成功落地首笔小微企业数据资产融资业务], *Sznews.com* [深圳新闻网], April 6, 20, [www.sznews.com/news/content/2023-04/06/content\\_30160729.htm](http://www.sznews.com/news/content/2023-04/06/content_30160729.htm)

with determination of data rights, data quality assessment, asset evaluation, and security and compliance assessment, played a crucial role in facilitating the deal while the exchange reviewed data products' compliance and the provider's qualifications.

Specifically, Guangdong Guanghe Law Firm provided a legal opinion and due diligence report for the trading data products sold by Weiyan Tech, covering issues of regulatory compliance and security. Moreover, the Shenzhen Institute of Standards and Technology assessed the quality of the data and whether intellectual property rights were protected, while Open Islands Community took care of the asset assessment. Subsequently, the Shenzhen Data Exchange reviewed all certifications and assessments and carried out the trading.

Introducing high-quality data products can also inspire trust and attract more customers. Suishenxing, Shanghai city's official smart transport platform and a registered data provider on the Shanghai Data Exchange, listed a product based on its smart parking system and managed to sign a deal with buyers at the end of 2022. As Shanghai's government-owned transport platform, Suishenxing covers the data of more than 4,700 garages and 890,000 parking lots.<sup>95</sup> After deep data cleansing and processing, the smart parking system can now provide real-time, accurate and comprehensive parking information to online mapping platforms, navigation software providers, garages, and other businesses.

### **Playing matchmaker**

Data exchanges can also facilitate trust-building by acting as matchmakers for transactions that otherwise may not materialize. Shenzhen Power Supply Co., Ltd, a subsidiary of China Southern Power Grid, developed an electricity data product called "electricity data-based credit inquiry" which analyzes enterprise's electricity data, including the status electricity usage, bill payment, power consumption, and records of breach of contract. The aim of the product is to provide banks and other financial institutions with information they can use to investigate the operation and credit profile of enterprises, thus helping them decide whether to provide financing.

The product was in demand: Financial institutions such as Shenzhen Credit Inquiry Service Platform (Company), Bank of Ningbo Shenzhen Branch, Duxiaoman Financial, and government departments such as the MIIT bureau of Shenzhen Baoan District all purchased it via the Shenzhen Data Exchange. Baoan District has used the data to evaluate whether to grant companies the government High- and New-Technology Enterprise status, one of China's main tax incentives for innovation. The Bank of Ningbo approved a loan to an electronic device manufacturing company in Shenzhen based on the information and analysis obtained through the product.<sup>96</sup>

The Guiyang Big Data Exchange similarly introduced several data products and services. One of them was an electricity data product. Guizhou Power Grid Corporation, a subsidiary of China Southern Power Grid signed a contract in April 2023 with Zhong Ding Credit, a credit rating agency, to provide data and analysis on enterprises' power consumption, including month-on-month and year-on-year ratio of electricity usage in the past three year; Zhong Ding Credit uses the data better assess enterprises' operations as part of their credit assessment.<sup>97</sup>

---

<sup>95</sup> [https://www.sohu.com/a/616894659\\_121488176](https://www.sohu.com/a/616894659_121488176)

<sup>96</sup> "[www.gov.cn/xinwen/2023-02/07/content\\_5740426.htm](http://www.gov.cn/xinwen/2023-02/07/content_5740426.htm)

<sup>97</sup> <https://news.cnstock.com/industry,rddj-202306-5079128.htm>

In a similar case, Shanghai Municipal Electric Power Company, a subsidiary of State Grid Corporation of China registered on the Shanghai Data Exchange and signed its first deal with ICBC Shanghai Branch for an electricity data-based product called Enterprises Electricity Smart Creative. The product is helping the bank determine whether issuing loans to enterprises based on their track record with electricity, including consumption, electricity usage behaviors, bill payment, and future projections.<sup>98</sup>

### Personal data trading

At present, personal information is effectively excluded from trading on data exchanges in China, although no local by-laws or regulations prohibit personal data trading. The most obvious explanation is that since the passage of the PIPL, legally permissible trading of personal data is extremely complicated and troublesome *de facto*. Effective and informed consent from every data subject is needed prior to any transfer of personal data among more than one entity.<sup>99</sup> The principles of legality, legitimacy, and avoiding any potential negative impact on individual rights and interests have concrete implications for data trading. Ensuring that data subjects' rights—such as modification, consent withdrawal, and deletion—are upheld can mean constant uncertainty for traders.<sup>100</sup> These legal restrictions have discouraged data exchanges from carrying out personal data trading.

The Guiyang Big Data Exchange, however, seems undeterred and became the first to perform personal data trading. Based on PET and other digital technologies, the recruiting platform Haohuo (Guizhou) desensitized and processed the resumes of job seekers as data products, such that any personally identifiable information would be hidden from users. The resume data product was listed on the Guiyang Big Data Exchange, which evaluated and assigned a reference price to the data product based on its self-developed pricing tool. A third-party service provider, a law firm in Guizhou, provided legal opinions. Individuals whose resumes are traded receive their revenue share from Haohuo.<sup>101</sup>

To the best of the authors' knowledge this is an isolated case, and it is hard to determine whether more such personal data products will be traded on China's data exchanges in the future. Importantly, the concept of privacy, which originated in liberal-democratic, rule-of-law polities, cannot be found in Chinese law where the very notion of individual right is absent.<sup>102</sup> Given that the PIPL does not impose any constraints on the ability of state organs to harvest citizens' data, the distinction between public and personal data is going to remain muddy. For example, the Shenzhen Special Economic Zone defines public data as information "generated and processed by public management and service agencies while managing or serving the public".<sup>103</sup> This effectively means that personal information may be embedded in a public data product and traded on Shenzhen's data exchange, with or without the data subjects' consent, and most definitely without any form of compensation."

### Takeaways

---

<sup>98</sup> <https://finance.sina.com.cn/jjxw/2021-11-25/doc-ikyakumx0179544.shtml>

<sup>99</sup> See Article 13- 32, Chapter II, Personal Information Handling Rules, Personal Information Protection Law in China.

<sup>100</sup> See footnote 82.

<sup>101</sup> [http://gz.news.cn/2023-05/11/c\\_1129605368.htm](http://gz.news.cn/2023-05/11/c_1129605368.htm)

<sup>102</sup> Creemers, Rogier. 2022. China's emerging data protection framework, *Journal of Cybersecurity*, Volume 8, Issue 1, published: August 24

<sup>103</sup> Shenzhen Special Economic Zone Data Regulation

This and the previous chapter have demonstrated how China's data exchanges 2.0—the Shenzhen Data Exchange, Guiyang Big Data Exchange, Beijing International Data Exchange, and Shanghai Data Exchange—have undertaken significant exploration and experimentation to solve thorny issues in China's data trading market, including data ownership, data pricing, and trust-building mechanisms. Crucially, the role of these exchanges is changing as their business models evolve from simple intermediaries to more sophisticated service providers and ecosystem orchestrators promoting data transactions. The Data Twenty Measures helped by casting off restrictions on data trading via official data exchanges, encouraging more market actors to make use of these platforms.

At the same time, data trading activity through data exchanges remains limited, and more progress is needed for China to fulfil its objective of creating a unified, efficient, and vibrant market of data factors fit for Chinese leaders' vision of dual circulation. So far, the government's direct support and coordination have played a significant role in getting the new generation of data marketplaces off the ground. Trust-building is not easy and takes time. The seemingly successful cases seen on the main data exchanges can be attributed to a large extent to government-led coordination among state-owned or state-linked participants. This makes it hard to determine the extent to which participation in the ecosystem is voluntary. For example, the data asset-based, CNY 10 million credit line to Shenzhen Weiyan Tech was instructed by the Shenzhen Municipal Government and the city's financial supervision agencies.<sup>104</sup>

Many of the deals in Guiyang, Beijing, Shanghai highlighted in this paper followed the same model and approach. China's data exchanges will surely face the challenge of keeping the momentum and enable data trading to transition towards a sustainable, market-oriented development model in the years to come, once the government's role has been scaled down—in fact, one wonders whether that is even the long-term vision.

## **Emerging trends: AI, cross-border transfers, and public data trading**

Besides remedying basic problems that hindered China's data circulation in the early days, Chinese data exchanges embody some broader and emerging trends which are also relevant for other data valuation and trading efforts around the world. This chapter will present three cases studies to respectively demonstrate the rapid growth of AI training data trading in China, the piloting of cross-border data trading, as well as some initial exploration of public (government) data trading. The analysis will examine both some promising areas of innovation and progress as well as some problems which accompany the development of data marketplaces in these areas.

### **Case Study 1: Trading of AI training data products on the Shanxi Data Exchange and Beijing International Data Exchange**

Data annotation and labeling is a rapidly growing market around the world. Per one estimate, this market will grow at a compound annual growth rate (CAGR) of 33.2 percent, reaching US\$ 3.6 billion by 2027 from US\$ 0.8 billion in 2022 and with the Asia Pacific region registering the highest CAGR.<sup>105</sup> More conservative projections suggest a CAGR of 26.5 percent from 2023-2030.<sup>106</sup> With AI gaining significant

---

<sup>104</sup> <https://bank.jrj.com.cn/2023/04/25095937503849.shtml>

<sup>105</sup> <https://www.researchandmarkets.com/reports/5744079>

<sup>106</sup> <https://www.grandviewresearch.com/industry-analysis/data-annotation-tools-market>



attention and investment globally, especially since ChatGPT was released in 2022, the AI training data market is booming and hit US\$ 1.7 billion in value in 2022.<sup>107</sup>

China's AI training data collection and annotation market began to emerge around 2010, in tandem with the rise of AI companies. Both specialized enterprises and large internet platforms began to enter the market. Per one estimate, the market value of China's basic data services for AI was CNY 4.5 billion in 2022 and will reach CNY 13-16 billion by 2027.<sup>108</sup> The approaches for data collection and annotation evolved from crowdsourcing and self-collection and annotation at the initial stage to a combining model, at which internet platforms or data companies build AI data annotation bases to collect and annotate data in an effective way and in a large scale. At present, AI data annotation in China remains a labor-intensive industry which mainly relies on manual annotation.<sup>109</sup>

Baidu, one of the leading AI companies in China, initially built up its own data annotation team to support AI training and development. More recently, growing demand for a wide array of use cases for AI training data, combined with stricter legal and regulatory requirements around data security, personal information protection, quality and efficiency has pushed the company to partner up with local governments and enterprises. In 2018, Baidu reached a deal with the Shanxi provincial government to co-establish an AI data annotation base. The largest of its kind nationwide, the base employed 5000 data annotators at 53 data annotation companies as of May 2022. Use cases in high demand include autonomous driving, voice and facial recognition, and mapping.

The Shanxi Data Trading Platform is exclusively focused on providing AI data collection and annotation products and services in the form of datasets, indexes, and API services for use cases such as voice recognition, facial recognition, and mapping data, in the areas of autonomous driving, smart terminal, smart security, smart retail, finance, healthcare, manufacturing, education, and translation services.<sup>110</sup> Cooperating with the Baidu AI Data Annotation Base allows the Shanxi Data Trading Platform to secure a stable supply of AI data products and services. In turn, the market demand from platform provides a useful feedback loop for Baidu to develop new types of AI data products and services.

As of this writing, 381 data products were listed on the Shanxi Data Exchange Platform, including 184 datasets and 197 APIs. Among these, 261 products were AI-related.<sup>111</sup> For example, there was a dataset called "human-vehicle interaction dialogue in Cantonese, Sichuanese dialect, and Mandarin with a Taiwanese accent," providing input for large language model training in Chinese with different dialects and accents. It contained records provided by a total of 5,300 individuals, with a man women ratio of 1:1 and an age range between 20 and 50. Interested buyers should contact the provider to negotiate the price.<sup>112</sup>

---

<sup>107</sup> <https://www.grandviewresearch.com/industry-analysis/ai-training-dataset-market>

<sup>108</sup> Deloitte. 2023. AI Basic Data Service White Paper, March.

<sup>109</sup> <https://www.semafor.com/article/03/02/2023/the-hidden-workers-in-china-influencing-ai-like-chatgpt>;  
<https://restofworld.org/2023/china-ai-student-labor/>

<sup>110</sup> See the website of the Shanxi Data Trading Platform at: <http://106.13.54.96/site/about/index>

<sup>111</sup> See the website of the Shanxi Data Trading Platform at:  
<http://106.13.54.96/datahub/tradepage/mall/list?publicMethod=0>

<sup>112</sup> See the website of the Shanxi Data Trading Platform at:  
<http://106.13.54.96/datahub/tradepage/mall/intr?id=1166>

A typical example of a popular API on the platform was a product called “faces comparison”. It could compare two faces and assign a similarity score for use cases such as person and ID verification and user verification. The product supported four types of images, including regular photos, photos of Chinese ID cards and other identification photos, and photos with grid lines. Users were charged between CNY 3000 and CNY 19000 yuan monthly, depending on the queries per second (QPS).<sup>113</sup>

In the future, Shanxi Data Exchange has big ambitions to become the biggest marketplace for AI data products in China and a one-stop “data factory” for data collection and annotation to boost industries from manufacturing to healthcare. Looking ahead, it is likely that more AI training data will be collected, annotated, and offered to companies such as Baidu through data exchanges like the Shanxi Data Exchange Platform. Other data exchanges in China are increasingly catching up and listing their own AI training data products to ride the wave of AI development.

At the same time, judged by the description of the listed AI data products, the Shanxi Data Exchange Platform still faces issues such as immature technology, a limited range and quantity of data sources, and difficulties in ensuring personal information protection. Such problems risk lowering the quality of data products and discouraging potential buyers.

The Beijing International Data Exchange also listed several AI training data products. Beijing Speechocean is an AI data resource and service provider. Holomatic is a startup company focused on providing autonomous driving solutions based on AI technology. In February 2022, the two companies signed a deal on AI algorithms training data at the Beijing International Data Exchange. Under the terms of the agreement, Holomatic engineers drove cars and collected real-world data including road conditions, traffic signals and signs, vehicles, pedestrians, and weather conditions. Holomatic performed initial data cleaning and processing and took care of desensitization and compliance inspections before transferring the data to Speechocean for annotation. Speechocean then returned more than 100,000 frames of annotated captured video data to Holomatic for AI algorithms training.<sup>114</sup>

In this case, the Beijing exchange was instrumental to the two companies reaching a deal. It first acted as a go-between and matchmaker between Holomatic and Speechocean, both partner enterprises of the exchange. It then assisted with reviewing compliance of the data trading entities, the data sources, the product as well as the use case of the data.

### **Case Study 2: Cross-Border Data Trading on the Shenzhen Data Exchange**

Currently, only a few major data exchanges in big cities in China, including Shenzhen, Shanghai, Beijing, offer cross-border data trading services. Among them, only the Shenzhen Data Exchange has trialed actual cross-border data (export) trading, while the Shanghai Data Exchange only provides data import services at its International Data Board. The Beijing International Data Exchange instead developed a data hosting service platform which supports enterprises’ cross-border data flows and started offering data hosting and desensitizing services to multinational corporations operating in China.<sup>115</sup>

The policy environment, insufficient administrative and technical capacity for compliance and security inspection, and strict data security review requirements are the major factors that have held cross-

---

<sup>113</sup> A measure of the amount of search traffic of an information-retrieval system.

<sup>114</sup> <https://finance.sina.cn/chanjing/gsxw/2022-11-28/detail-imqmmthc6229849.d.html?from=wap>

<sup>115</sup> <http://www.bjchy.gov.cn/lqjs/lqdt/4028805a811a47da01811f14f64e046b.html>

border data trading back. One important reason behind Shenzhen's advances with cross-border data transmissions is that the city has secured policy support from the central government for trialling cross-border data trading. In January 2022, the NDRC and the Ministry of Commerce (MOFCOM) jointly issued the "Opinions on Several Special Measures to Relax Market Access for Shenzhen Building a Pilot Demonstration Zone for Socialism with Chinese Characteristics." Article 2 of the first section of the Opinions encourages Shenzhen to experiment with data cross-border (export) security management mechanisms and explore to establishment of offshore data trading platforms.

Experiments around international data ports are also underway in the Shanghai Ligang District, Hainan Free Trade Port, and Guangzhou—the latter focused on economic integration within the Greater Bay Area.<sup>116</sup> However, not much has happened since MOFCOM in 2020 called on free trade zones in these and other localities to stimulate cross-border data trading.<sup>117</sup> Most of these pilots so far have focused on building supporting digital infrastructure, such as submarine optical fiber cable, data centers, and industrial parks for big data. The fact that the expected gateways for "safe and orderly" data transmissions across borders have not materialized may have persuaded the central government to bet on Shenzhen.

Shenzhen borders Hong Kong and has traditionally been a leading experimental zone for China's market-oriented reform and opening-up policy since the end of the 1970s. It now shoulders the task of exploring arrangements and mechanisms for further relaxing market access and trade restrictions in the digital age, which also entails the trialling of cross-border data flows. This goal was also highlighted in the Data Regulations of the Shenzhen Special Economic Zone issued in July 2021, the first in China to propose a definition of data ownership rights.<sup>118</sup> In addition this regulatory foundation, the Shenzhen Data Exchange enjoys solid policy support from local governments. The Hong Kong-Shenzhen Innovation and Technology Cooperation Zone where the exchange is located has introduced policies to promote data resource sharing and technological cooperation on AI and data analysis between Hong Kong and Shenzhen. Among the 14 deals reached until the end of 2022, eight were data exports between Shenzhen and Hong Kong.<sup>119</sup>

As of March 2023, 16 cross-border deals had been closed through the exchange.<sup>120</sup> The platform had facilitated trading for a cumulative value of CNY 11 million yuan by November of the previous year.<sup>121</sup> The first deal, worth CNY 5 million, involved a foreign hedge fund and the domestic data provider ChinaScope. The buyer sought to purchase data from ChinaScope's flagship data product, Smartag news analysis engine, which uses a self-developed National Language Processing (NLP) algorithm to convert unstructured Chinese language news text into machine-readable metadata. The algorithm extracts, tags

---

<sup>116</sup> [www.sohu.com/a/652419269\\_121255906](http://www.sohu.com/a/652419269_121255906); Research on mechanisms of safe and orderly cross-border data flows at Hainan Free Trade Port 2021-2022; [http://hmo.gd.gov.cn/ns/content/post\\_3755560.html](http://hmo.gd.gov.cn/ns/content/post_3755560.html)

<sup>117</sup> <https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/China/Policy-Briefing-Cross-BorderDataTransfer.html>

<sup>118</sup> <https://www.chinalawtranslate.com/en/shenzhen-data-regs/>

<sup>119</sup> <https://m.mp.oeeee.com/a/BAAFRD000020220414671907.html>; <https://www.hkpc.org/en/about-us/background>

<sup>120</sup> [http://www.sz.gov.cn/cn/xxgk/zfxxgj/bmdt/content/post\\_10586630.html](http://www.sz.gov.cn/cn/xxgk/zfxxgj/bmdt/content/post_10586630.html)

<sup>121</sup> [https://finance.sina.com.cn/china/gncj/2022-12-03/doc-imqmmthc6944653.shtml?cre=tianyi&mod=pcfinf&loc=4&r=0&rfunc=18&tj=cxvertical\\_pc\\_fin&tr=12](https://finance.sina.com.cn/china/gncj/2022-12-03/doc-imqmmthc6944653.shtml?cre=tianyi&mod=pcfinf&loc=4&r=0&rfunc=18&tj=cxvertical_pc_fin&tr=12)

and classifies a wide range of information points from hundreds of Chinese media outlets in near real time, compiling sentiment indicators and events linked to Chinese companies to help users to stay ahead of market trends and risks.<sup>122</sup>

The quality assurance, compliance analysis and security audit performed by the Shenzhen Data Exchange are proving key in reaching data export deals. Compliance and security audit are especially important for both data providers and buyers, considering China's strict data protection, data security, and localization requirements. The process typically entails three steps. First, ChinaScope submitted basic information such as the name and description of its data product, information about itself and buyers, use cases, security of data, as well as the necessary approvals and certificates from relevant government agencies. Second, a law firm provided a legal opinion for compliance review, which paid attention to the parties' capacity to guarantee data security and comply with relevant laws and regulations on cross-border data transfer.

As a third step, the Shenzhen Data Exchange conducted its own compliance review to assess, *inter alia*: the legality, justifiability, and necessity of the transaction; the purpose and use cases; the amount, scope, category and sensitivity of the data being transferred as well as any potential risks the transfer could pose to China's national security, public interest, and the legitimate rights and interests individuals and organizations, in line with the requirements of the Data Security Law; and finally, the ability of the data processor to put in place the necessary governance and technological measures to prevent data leakage, damage and other risks.

As mentioned, other exchanges are more focused importing data into China. Most of the 28 data products listed on the International Board of the Shanghai Data Exchange since its inception in April 2023 target domestic enterprises that have a demand for importing data.<sup>123</sup> Typical providers include China National Publications Import and Export Company, which specializes in such datasets as statistical information, patents, and encyclopaedic resources. Another one, PatSnap Suzhou Co. Ltd., listed three data products which had undergone cleansing and verification and were based on a global patent database partly purchased from the United States Patent and Trademark Office and the European Patent Office.<sup>124</sup>

It is obvious that the laws and regulations issued over the past three to five years have paved the way for a more complete and predictable data governance regime. Chinese regulators have actively promoted and incentivized new mechanisms for outbound data transfers and have encouraged local government bureaucracies, companies, and data exchange platforms to carry out experiments. Beyond data exchanges, CAC bureaus in Beijing and Shanghai have begun promoting typical cases for data export security review to set examples for data export nationwide.<sup>125</sup> Data export trials in localities like Shenzhen are receiving policy support from both the central and the local governments.

The question on sustainability remains unclear, though. After the first few attempts and initial achievements driven by government coordination behind the scenes, data exchanges seem to be facing almost the same difficulties as before in promoting data transactions and exports. A more streamlined

---

<sup>122</sup> <https://www.chinascope.com/ai-news.html>; <https://m.mp.oeeee.com/a/BAAFRD000020220515683739.html>

<sup>123</sup> See International Board at Shanghai Data Exchange at: <https://nidts.chinadep.com/ep-hall>

<sup>124</sup> <https://web.shobserver.com/staticsg/res/html/web/newsDetail.html?id=615395>

<sup>125</sup> Ibid; <http://www.glo.com.cn/Content/2023/02-15/0845294502.html>

process in the CAC-mandated security reviews, further relaxation of the strict restrictions on data export, clearer definitions for key data categories such as important and core data, and more market-oriented approaches would be required for China to further promote cross-border data trading.

### **Case Study 3: Government/Public Data Trading on the Beijing International Data Exchange and the Hainan Supermarket for Data Products**

Public data has great economic potential, and the exploitation and utilization of public data can be a crucial component of unleashing the value of data elements around the world. Already ten years ago, it was estimated that the use of open data (not limit to government data) could unlock USD 3 trillion in economic value.<sup>126</sup> According to another estimate from 2018, data and data analytics could contribute USD 1.2 trillion a year to the public and social sector.<sup>127</sup> Based on the methodology and approach of a McKinsey report on big data from 2011,<sup>128</sup> the state-owned outlet People's Daily estimated the potential value of government data openness in China to reach between CNY 10 and 15 trillion in 2024.<sup>129</sup>

More than 80 percent of China's information and data resources have been said to be in the hands of government bureaucracies at all levels.<sup>130</sup> Then Chinese Premier Li Keqiang mentioned the figure in 2016 when trying to push the government to break "information silos (islands)" and optimize governmental services.<sup>131</sup> Since China unveiled its big data strategy in 2015, authorities at all levels have pushed for public data opening via government-run platforms. More than 670 local governments' open data platforms were online as of the end of 2021.<sup>132</sup> With more data being produced, collected, and analyzed by large digital platforms and their intricate ecosystems of apps, the situation today looks very different compared to almost a decade ago. Nevertheless, public data still accounts for a significant portion of China's data resources.

China's open data government platforms suffer from serious quality problems. Nearly 85 percent of government data collected and provided for public inquiry is incomplete.<sup>133</sup> Poor data quality, slow update frequency, and data inconsistency have rendered many platforms nearly useless for business and individual users. Most departments only started these platforms to satisfy superior organs. The lack of clear rules and governance mechanisms, poor data quality and standardization, and the absence of revenue-sharing arrangements have made many of them unattractive and inefficient. Another hurdle stems from concerns over data security and information protection, since public data may contain information that is critical for national security and public safety, business secrets, or personal information.

---

<sup>126</sup> Open data: Unlocking innovation and performance with liquid information.

<sup>127</sup> <https://www.mckinsey.com/industries/public-sector/our-insights/accelerating-data-and-analytics-transformations-in-the-public-sector>.

<sup>128</sup> The report estimated, base on the data on OECD-Europe public sector, that big data has the potential to save 15-20 percent of operating expenditure, reduce 30-40 percent of fraud and error in transfer payment, and increase 10-20 percent of tax collection. See: Big data: The next frontier for innovation, competition, and productivity.

<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>

<sup>129</sup> <http://cpc.people.com.cn/n/2014/0521/c83083-25044169.html>

<sup>130</sup> [https://www.gov.cn/xinwen/2016-05/13/content\\_5073036.htm](https://www.gov.cn/xinwen/2016-05/13/content_5073036.htm)

<sup>131</sup> [https://www.gov.cn/zhengce/2016-12/08/content\\_5145134.htm](https://www.gov.cn/zhengce/2016-12/08/content_5145134.htm)

<sup>132</sup> <https://www.secrss.com/articles/54338>

<sup>133</sup> <http://www.caict.ac.cn/kxyj/qwfb/bps/202304/P020230427572038320317.pdf>

Data exchanges could offer solutions to these problems by creating effective pathway for high quality public data circulation in China, provided that clear rules and standards are in place. The major exchanges have already developed the capacity to provide more stable, secure, and trustworthy environments for data trading, including access to professional third-party services and technological support such as PET. Recently, the Data Twenty Measures explicitly directed data exchanges to be public institutions serving the public interest. Consequently, some exchanges, like the Shanghai Data Exchange, are building platforms as quasi-public institutions acting as public infrastructure and providing services as non-profit organizations. These are typically invested by the government or state-owned assets and charge a reasonable price for the services provided.<sup>134</sup>

In Beijing, the government is seeking to entrust professional entities and data exchanges with managing government open data platforms. In the area of finance, the government authorized Beijing Financial Holdings Group (BFHG), a state-owned financial company established in 2018, to set up a Financial Public Data Zone and manage government financial data. It provides public data products and services in the form of credit inquiries, market access analyses, risk assessments, and analyses of enterprise competitiveness and credit records covering 3.4 billion data points across industries and businesses, judiciary, taxation, social security, housing provident funds, and real estate. The data come from 14 government agencies and 2.7 million market entities, and it updated daily, weekly, and monthly.<sup>135</sup>

In 2021 the BFHG initiated and registered Beijing International Data Exchange Co. Ltd., where it holds 65 percent of the shares.<sup>136</sup> The exchange has direct access to the zone, as well as to the Beijing Municipal Public Data Portal which is directly run by the Beijing Commission of Economy and Information Technology, the government agency that manages Beijing's public data and authorized the establishment of the zone. In other words, the Beijing International Data Exchange was entrusted to run the entire public data resources of the Beijing municipal government, and to turn the traditional government open data platform into a data marketplace.

Government/public data products and services account for nearly half of the 96 listings on the Beijing International Data Exchange.<sup>137</sup> Typical listings spans credit inquiry as well as business and industry information based on government data (See Appendix). Some interesting observations can be drawn from examining of the public data products and services listed on the exchange.

First, most constitute unstructured statistical data—only three are processed index products with quantified data assessments. The quality of data products and services is debatable, and some are also available on the Beijing's government open data platform where a much greater variety of public data is available and free to access.<sup>138</sup>

---

<sup>134</sup> [https://www.chinadep.com/bulletin/notices/CTC\\_20230804153158042745](https://www.chinadep.com/bulletin/notices/CTC_20230804153158042745). However, the Shanghai Data Exchange does not yet offer public data products and services comparable to those that are traded in Beijing. <https://nidts.chinadep.com/ep-hall>.

<sup>135</sup> [http://www.jjckb.cn/2023-01/11/c\\_1310689641.htm](http://www.jjckb.cn/2023-01/11/c_1310689641.htm)

<sup>136</sup> <https://www.bjidex.com/infoDetail/NZzR4kKHlt>. The two entities also share the same person as the chairman of the board, <https://www.bfhg.com.cn/html/gltld/>

<sup>137</sup> Calculated based on the online data trading market at Beijing International Data Exchange: <https://webs.bjidex.com/sys-bsc-home/#/bscConsole/tradingMarket>

<sup>138</sup> See [data.beijing.gov.cn](http://data.beijing.gov.cn) for the platform.

Second, it is apparent that the purpose of the Beijing International Data Exchange is still limited to a go-between for data providers and buyers. A subscription is required to view product and service details and other information such as terms of use and user guidelines, and delivery takes place off the exchange's platform. Moreover, there is not mention of prices, which instead are supposed to be set during in-person negotiation.

Third, half of the listed data products are credit inquiry data services provided by Beijing Financial Big Data Company via its own platform "Jingyun Credit Inquiry," which provides credit inquiry services to all agencies in Beijing involved in offering financing guarantees. Based on the rich data resources of the Financial Public Data Zone, Jingyun Credit Inquiry provides services such credit investigation, credit reports, risk warning, and post financing/guarantee early warning.<sup>139</sup> Yet a look at Jingyun Credit Inquiry's website suggests it has much richer data resources than those it makes available for trading through the Beijing International Data Exchange.

In contrast to the authorized operation model at the Beijing International Data Exchange, where a subsidiary company was entrusted with running all public data products, the government-run model in Hainan explores an ecosystem with multiple actors and data developers.

The Hainan Supermarket for Data Products is a unique data exchange platform run directly by the provincial Big Data Administration of the Hainan government, with technology support from China Telecom Hainan Company and its cloud subsidiary Tianyi Cloud, one of China's largest cloud service providers. It combines the functions of an online data marketplace and a government open data platform.<sup>140</sup> To overcome issues such as data security, privacy protection, and compliance, the Hainan Big Data Administration invited data developing companies to the platform and develop products and services based on data resources provided by the Hainan government. These data products and services are listed and traded at the Hainan Supermarket for Data Products.<sup>141</sup>

The Hainan Supermarket for Data Products features a wide variety of data products and services and seems more active than the Beijing exchange. In total, 831 provincial public datasets (APIs) were provided by Hainan government agencies at all levels and 696 national public datasets (APIs) were provided by government-affiliated companies, such as China Economic Information Service of Xinhua News Agency, privately owned credit inquiry companies, and business data platforms like Tianyancha.<sup>142</sup> Based on these data resources, the invited companies developed and commercialized 1070 data products and services through the Hainan Supermarket for Data Products.<sup>143</sup>

The responsibility of guaranteeing data security, privacy protection and compliance was entrusted to the technology partner, Tianyi Cloud, which helped build a data management system in which public data was desensitized. Using PET, secure multi-party computation, and federated learning,<sup>144</sup> Tianyi Cloud guarantees that public data is traded in a secure, compliant and privacy-preserving manner. In this way, government/public data resources are flowing orderly in the forms of reliable and tradable products and

---

<sup>139</sup> <https://www.bjzhengxin.com.cn/aboutJy?num=>

<sup>140</sup> <https://www.datadex.cn/home>

<sup>141</sup> <https://www.govmade.cn/viewpoint/20211124/648121624679153664.html>

<sup>142</sup> <https://www.datadex.cn/resourceList/index>

<sup>143</sup> <https://www.datadex.cn/app/dataMarket>

<sup>144</sup> <https://www.ctyun.cn/cases/596200642071100416>

services to entities that need them, while the underlying data resources remain safely in the hands of the government-owned platform.

### **Takeaways from the case studies: a state-led data economy**

China's data exchanges have become avenues for regulatory, technological, and institutional experimentation. As a range of industries are heavily invested in AI development and applications, AI training datasets are now among the main products on the exchanges. China is also seeking to promote cross-border digital trade by testing and trialling the controlled export of data, including through the young CAC-led mechanism of security review. Additionally, authorities are trying to establish mechanisms for trading government/public data, either by authorizing government-affiliated enterprises to run government data trading through data exchanges, or by having the government directly in charge.

The fact that most data exchanges are controlled by state-owned assets and the deep involvement of the state in their operations show one clear feature of Chinese data exchanges: that of being government-centric. Judging from the business practices of major data exchanges studied in this paper, public-private partnerships are playing a significant role in promoting data transactions. The three case studies presented in the previous chapter all demonstrate how China is seeking to create a state-led data market.

More specifically, the government performs a coordinating function and sometimes introduces data providers and buyers to data exchanges, as demonstrated by the cases of Beijing and Hainan as well as data export trials in Shenzhen. It is also clear that public-private partnerships between local governments and digital platform companies are behind some of the data products and services being traded on the exchanges, as seen with AI training products on the Shanxi Data Exchange. Even beyond this particular case, state-owned enterprises in the fields of finance, electricity, telecom, companies with strong government connections, and tech companies dominate data providers and buyers on China's main data exchanges are.

Overall, the role of market forces in the development and operations of data exchanges has not yet been fully demonstrated. Top-level policy planning documents state that "liberating digital productive forces" will require the government to direct the market more vigorously.<sup>145</sup> Although more private companies are participating in China's main data exchanges, it remains to be seen whether their involvement and initiative will be allowed to grow to such a level as to allow China's data trading market to develop in a more market-oriented direction.

### **Conclusion: China's data market in the global context**

The establishment and practice of data marketplaces in China represents quite a unique phenomenon on the global stage. Indeed, the rise of large digital platform businesses, such as Google and Meta, has coincided with a concentration of data power in the hands of few private enterprises. A wider ecosystem of data brokerage firms has developed and thrived around these platforms, particularly in the United States.<sup>146</sup> Brokers such as Acxiom, Data Axle and LiveRamp, large credit reporting agencies like Equifax, more traditional data analytics companies such as LexisNexis and CoreLogic, as well as large

---

<sup>145</sup> <https://digichina.stanford.edu/work/translation-14th-five-year-plan-for-national-informatization-dec-2021/>

<sup>146</sup> <https://www.wired.com/story/opinion-data-brokers-are-a-threat-to-democracy/>



technology companies and cloud service providers such as Oracle are regarded as some of main data providers and brokers in the world. These players do not make use of any data intermediaries, instead keeping control over data sales and trading in their own hands.

Countries like India, Colombia and Japan have explored the concept of data exchanges. For example, India's Data Empowerment and Protection Architecture (DEPA) introduced techno-legal structures, including so-called consent managers, to facilitate consent-based data sharing between data controllers and users in the financial and health sectors.<sup>147</sup> The World Economic Forum has elevated the issue in importance by creating the Data for Common Purpose Initiative, a multistakeholder community trying to unlock the value of data.<sup>148</sup> However, the next most ambitious endeavor underway outside China is arguably in the European Union, where the Data Governance Act and the yet-to-be adopted Data Act aim to set the structures and rules for trusted data sharing, access and use within European data spaces.<sup>149</sup>

To our knowledge, however, none of these experiments go as far as China's efforts at creating institutional marketplaces for digital asset trading. China was a clear first mover in elevating big data to national strategic priority, and subsequently designating data as a factor of production. Such actions reflect a uniquely state-driven approach to data governance and managing the digital economy which also characterizes the design and practices of Chinese data exchanges. The first generation of data exchanges were partly driven by market forces operating in a near-total absence of regulation. The ensuing economic inefficiency and public policy dilemmas persuaded the central government that the market-driven model should give way to stronger government intervention and oversight. Beijing's push to have the state direct data circulation, in line with Xi Jinping's lofty plans for a 'Digital China', provides the momentum behind the ongoing reform.

These developments will carry global implications beyond offering lessons for other jurisdictions, also considering that Chinese authorities are exploring how some of these marketplaces could serve as gateways for cross-border data transmissions.<sup>150</sup> Given that transnational data flows are considerably more restricted and less transparent at China's borders than they are at those of OECD economies, the extent to which the country will seek to integrate its nascent data market with those of its trading partners, or prioritize domestic circulation in the name of national security, remains to be seen. Foreign companies that conduct business in China, meanwhile, will want to monitor the ongoing institutionalization of domestic data trading to be prepared for any potential opportunities. One area to watch is the emerging experimentation around the trading of AI training data products.

There are encouraging signs that the balance may tilt towards a partial relaxation of China's strict data localization requirements, which may signal real concerns in some parts of the government that excessive securitization is undermining efforts to unleash the economic potential of data. In a major departure from its security-first approach, the CAC in October 2023 released draft rules in which it proposed to ease the burden of data classification on business.<sup>151</sup> Among other measures, the rules

---

<sup>147</sup> <https://www.orfonline.org/research/data-empowerment-and-protection-architecture-concept-and-assessment/>

<sup>148</sup> <https://www.weforum.org/agenda/2022/01/data-trading-stock-exchange/>

<sup>149</sup> <https://digital-strategy.ec.europa.eu/en/policies/data-act>

<sup>150</sup> <https://merics.org/de/kommentar/beijings-watchful-eye-all-data-flowing-and-out-china>

<sup>151</sup> <https://archive.ph/I9VG2>

would exempt various cross-border data export scenarios, such as academic collaborations and manufacturing, from the mandatory security review, and authorize companies to transfer data abroad when it has not been classified as “important” by regulators. The idea would be to let business, not regulators, decide when cross-border data flows are necessary for their global operations.<sup>152</sup> This followed the State Council’s call for free data flows in a set of measures released in August to restore business confidence, amid sluggish growth and industry’s frustration with tightening party-state control over the economy.<sup>153</sup>

China has been remarkably innovative at opening up select government and public data for trading, to the benefit of companies that could develop relevant products and services. Still, compared to government open data in developed countries such as the United States, Canada and Germany, its government open data platforms have faced significant challenges such as low data quality and slow update frequency, to the point of making some of the data almost useless or even undesirable. The Chinese government has been trying to encourage the development of data products and services based on public data for quite a long time and through a number of policies, some which predated efforts at creating a national data market. This is where data exchanges came in, trying to use their established mechanisms and arrangements to enable public data to be traded in a more open, efficient, and accountable manner.

In an international context, this paper has demonstrated that although China is ahead of any other country in developing a data trading market, this fact alone is no guarantee for the quality of data products and services being traded at the data exchanges. China has still a long way to go before it can foster an efficient, vibrant, and well governed data market. For real demand to emerge, more high-quality data products and services will need to be issued in the market. One way to achieve this goal may be for more market-oriented participants and private companies to be engaged in the process. This, however, would require some toning down of the government-centric features that still characterize China’s model of data circulation.

## Appendix

**Table1. Data exchanges and trading platforms in mainland China**

Name	Name in Chinese	Founding Year	Location	Note
Zhongguancun Shuhai Big Data Trading Service Platform	中关村数海大数据交易服务平台	2014	Beijing	First data trading platform in China; not active
Beijing Big Data Trading Service Platform	北京大数据交易服务平台	2014	Beijing	Not active
Guiyang Big Data Exchange	贵阳大数据交易所	2015	Guizhou, Guiyang	First data exchange in China
Chongqing Big Data Trading Platform	重庆大数据交易平台	2015	Chongqing	Not active
Wuhan East Lake Trading Center for Big Data	武汉东湖大数据交易中心	2015	Wuhan, Hubei	Not active

<sup>152</sup> <https://www.21jingji.com/article/20230929/4b3de75d9185ebfc77ce1c22aff5e36e.html>

<sup>153</sup> [https://english.www.gov.cn/policies/policywatch/202308/14/content\\_WS64d9680bc6d0868f4e8de85f.html](https://english.www.gov.cn/policies/policywatch/202308/14/content_WS64d9680bc6d0868f4e8de85f.html)

Draft paper for IARIW-CIGI Conference on the Valuation of Data

Wuhan Changjiang Big Data Trading Center	武汉长江大数据交易中心	2015	Wuhan, Hubei	Evolved into Changjiang Data Exchange
Central China Big Data Exchange	华中大数据交易所	2015	Wuhan, Hubei	Not active
East China Jiangsu Big Data Exchange Center	华东江苏大数据交易中心	2015	Yancheng, Jiangsu	
Hebei Big Data Trading Center	河北大数据交易中心	2015	Chengde, Hebei	Not active
Transportation Big Data Trading Platform	交通大数据交易平台	2015	Shenzhen, Guangdong	Not active
Hangzhou Qiantang Big Data Trading Center	杭州钱塘大数据交易中心	2015	Hangzhou, Zhejiang	Focusing on industrial data
Xixian New District Big Data Exchange	西咸新区大数据交易所	2015	Xi'an, Shaanxi	Not active
Shanghai Data Exchange Center	上海数据交易中心	2016	Shanghai	Evolved into Shanghai Data Exchange in 2021
Zhejiang Big Data Exchange Center	浙江大数据交易中心	2016	Hangzhou, Zhejiang	
Harbin Data Trading Center	哈尔滨数据交易中心	2016	Harbin, Heilongjiang	Not active
Guangzhou Data Trading Service Platform	广州数据交易服务平台	2016	Guangzhou, Guangdong	Not active
Southern Big Data Trading Center	南方大数据交易中心	2016	Shenzhen, Guangdong	
Silk Road Big Data Trading Center	丝路辉煌大数据交易中心	2016	Lanzhou, Gansu	Not active
Qingdao Big Data Trading Center	青岛大数据交易中心	2017	Qingdao, Shandong	
Henan Pingyuan Big Data Trading Center	河南平原大数据交易中心	2017	Xinxiang, Henan	Not active
Henan Zhongyuan Big Data Trading Center	河南中原大数据交易中心	2017	Zhengzhou, Henan	Not active
Jilin Northeast Asian Big Data Trading Service Center	吉林省东北亚大数据交易服务中心	2018	Changchun, Jilin	Not active
Shandong Data Exchange Platform	山东数据交易平台	2019	Jinan, Shandong	
Beibu Gulf Big Data Trading Center	北部湾大数据交易中心	2020	Nanning, Guangxi	
Shanxi Data Exchange Platform	山西数据交易平台	2020	Taiyuan, Shanxi	Featuring AI data collection and annotation
Zhongguancun Medicine and Health Big Data Trading Platform	中关村医药健康大数据交易平台	2020	Beijing	Specialized data trading platform
Shanghai Data Exchange	上海数据交易所	2021	Shanghai	
Beijing International Data Exchange	北京国际大数据交易所	2021	Beijing	
Western China Data Exchange	西部数据交易中心	2021	Chongqing	

North Big Data Exchange Center	北方大数据交易中心	2021	Tianjin	
Hefei Data Factor Circulation Platform	合肥数据要素流通平台	2021	Hefei, Anhui	
Yangtze River Delta Data Factor Circulation Service Platform	长三角数据要素流通服务平台	2021	Suzhou, Jiangsu	
Hainan Supermarket for Data Products	海南省数据产品超市	2021	Haikou, Hainan	
South China International Data Exchange Co. Ltd	华南(广东)国际数据交易公司	2021	Foshan, Guangdong	
Shenzhen Data Exchange	深圳数据交易所	2022	Shenzhen, Guangdong	
De Yang Data Exchange	德阳数据交易中心	2022	Deyang, Sichuan	
Zhengzhou Data Exchange Center	郑州数据交易中心	2022	Zhengzhou, Henan	
Fuzhou Big Data Exchange	福建大数据交易所	2022	Fuzhou, Fujian	
Hunan Big Data Exchange	湖南大数据交易所	2022	Changsha, Hunan	
Qingdao Oceanic Data Exchange Platform	青岛海洋数据交易平台	2022	Qingdao, Shandong	
Wuxi Big Data Exchange Platform	无锡大数据交易平台	2022	Wuxi, Jiangsu	
Guangzhou Data Exchange	广州数据交易所	2022	Guangzhou, Guangdong	

*Source: Authors' research.*

**Table 2. Five major data exchanges in mainland China**

Name	Organization Type	Business model	Main products	Total Trading Volume
Guiyang Big Data Exchange	State-owned	Data value-added services	Data products and services, algorithmic tools, and resources	CNY 1.4 billion as of July 2023 <sup>154</sup>
Shenzhen Data Exchange	State-owned	Data value-added services	Data products, services, and tools	CNY 1.8 billion as of March 31, 2023, tops other data exchanges <sup>155</sup>
Shanghai Data Exchange	State-owned assets holding	Quasi-public service institution, data service fee	Data sets, data services	CNY 0.1 billion as of December 2022 <sup>156</sup>
Beijing International Data Exchange	State-owned assets holding	Data value-added services	Data products including data sets, API, reports, and data services	Not available

<sup>154</sup> <https://www.gzdex.com.cn/>

<sup>155</sup> [http://szzf.gd.gov.cn/2022szzfjfh/szzf2022/content/post\\_4174214.html](http://szzf.gd.gov.cn/2022szzfjfh/szzf2022/content/post_4174214.html);

[http://www.sz.gov.cn/cn/xxgk/zfxxgj/bmdt/content/post\\_10586630.html](http://www.sz.gov.cn/cn/xxgk/zfxxgj/bmdt/content/post_10586630.html)

<sup>156</sup> <https://www.stcn.com/article/detail/767816.html>

Guangzhou Data Exchange	State-owned assets holding	Data value-added services	Data products and services, data resources, data assets	CNY 1 billion as of May 2023 <sup>157</sup>
-------------------------	----------------------------	---------------------------	---	---

**Note: Except for the Guangzhou Data Exchange, all the exchanges listed here offer a public online platform for data trading.**

**Source: Authors' research.**

**Table 3. Other 12 active data exchange platforms**

Name	Organization Type	Business model	Main products	Total Trading Volume
North Big Data Exchange Center	Mixed ownership with state-owned assets' shares	Data value-added services	Data products and service.	CNY 0.15 billion (aspirational target as of May 2023) <sup>158</sup>
East China Jiangsu Big Data Exchange Center <sup>159</sup>	Joint Stock Company	Membership annual fee	Data products and services	Not available
Zhengzhou Data Exchange Center	State-owned assets holding	Quasi-public service institution providing data value-added services	Data products and services	CNY 0.1 billion as of June 2023 <sup>160</sup>
Western China Data Exchange	100 percent state assets owned	Data value-added services	Data products	CNY 0.1 billion as of January 2023 <sup>161</sup>
Changjiang Data Exchange	State-owned assets holding	Data trading and renting services; membership fee <sup>162</sup>	Data products	Not available
Zhejiang Big Data Exchange Center	State-owned assets holding	Commission fee, membership fee, and data service fee	Data sets, API, reports, algorithmic models, data services	Not available
De Yang Data Exchange	State-owned assets holding	Data trading services	Data products and services	CNY 23.8 million as of July 2023
Shanxi Data Exchange Platform	Joint platform run by Shanxi government and Baidu	Data trading services and data value-added services	AI data sets, API, index	CNY 50 million as of March 2021 <sup>163</sup> (latest available figure)
Shandong Data Exchange Platform	State-owned provincial data service platform	Data trading services and data value-added services	Data sets, reports, applications, API, privacy-enhancing computing, data services	Not available; 2022 revenues were CNY 14.4 million; 2022 net profits CNY 1.2 million

<sup>157</sup> [http://zfsq.gd.gov.cn/xxfb/ywsd/content/post\\_4178925.html](http://zfsq.gd.gov.cn/xxfb/ywsd/content/post_4178925.html)

<sup>158</sup> <https://www.datadmz.com/zh/news/bei-fang-da-shu-ju-jiao-yi-zhong-xin-zheng-shi-jie-pai-cheng-li>

<sup>159</sup> Its website [www.Hddatapay.com](http://www.Hddatapay.com) can not be reached.

<sup>160</sup> <https://www.zzbdex.com/newsDetail?id=ffc074012adc411b89ec77643c5f394a&category=0>

<sup>161</sup> [http://www.cqjb.gov.cn/bm/qdsjfzj\\_71933/zwx\\_73796/dt/202301/t20230106\\_11464299.html](http://www.cqjb.gov.cn/bm/qdsjfzj_71933/zwx_73796/dt/202301/t20230106_11464299.html)

<sup>162</sup> <https://www.cjdataex.cn/lmhy>

<sup>163</sup> [https://www.sohu.com/a/458108442\\_120214183](https://www.sohu.com/a/458108442_120214183)

Beibu Gulf Big Data Trading Center	State-owned assets holding	Data value-added services; authorized use of data or direct purchase	Data set, API, solutions	CNY 15 million as of 2020 (latest available figure)
Hefei Data Factor Circulation Platform	State-owned by Hefei Big Data Asset Operation Co. LTD.	Data value-added services	Data products (data sets, API, reports), services, and tools	CNY 41 million as of June 2023 <sup>164</sup>
Hainan Supermarket for Data Products	Run by Hainan provincial government	Platform for public data products	Data sets, API, reports, models, data services	CNY 0.4 billion as of July 2023

**Note:** Except for the East China Jiangsu Big Data Exchange Center, all the data exchanges listed here offer a public online platform for data trading.

Source: Authors' research.

**Table 4: List of government/public data products and services at the Beijing International Data Exchange**

	Name	Category		Name	Category
1	Information on dishonest enterprises committing grave illegalities	Credit inquiry	24	Court announcements	Credit inquiry
2	Legal persons in grave tax-related violation cases	Credit inquiry	25	Litigation information of enterprises	Credit inquiry
3	Information of notice of tax arrears	Credit inquiry	26	List of high-tech enterprises in Zhongguancun	Business & industry
4	Information on tax-related administrative punishments	Credit inquiry	27	Enterprises with food business license	Business & industry
5	Information on grave tax-related violation cases	Credit inquiry	28	Outbound investment of enterprises	Business & industry
6	List of dishonest enterprises in customs	Credit inquiry	29	Information verification of small and micro businesses	Business & industry
7	List of advanced certified enterprises authorized by customs	Credit inquiry	30	Change records of enterprises	Business & industry
8	Seriously dishonest enterprises in statistics (list of punishments since July 2019)	Credit inquiry	31	Installment payments of investors	Business & industry
9	Information on administrative punishments for enterprises' housing provident funds	Credit inquiry	32	Information on investors	Business & industry
10	Information on enterprises with non-performing housing provident funds	Credit inquiry	33	Customized data mining of newly registered enterprises	Business & industry
11	Push notification of enterprises' early-warning information	Credit inquiry	34	Human resource competitiveness index of technological enterprises	Business & industry
12	Notice of tax arrears	Credit inquiry	35	Index of enterprises' profitability	Business & industry
13	Information on serious violations	Credit inquiry	36	Information on quality management awards in Beijing	Business & industry

<sup>164</sup> [https://www.sohu.com/a/691807768\\_120133855](https://www.sohu.com/a/691807768_120133855)

14	Information verification of enterprises' early warning	Credit inquiry	37	Index of enterprises' overall competitiveness	Business & industry
15	Information on administrative punishments in water supplies	Credit inquiry	38	High-tech enterprises information retrieval	Business & industry
16	Notice of court sessions	Credit inquiry	39	Information on municipal government procurement contracts	Business & industry
17	Tax-related violations - Beijing data set	Credit inquiry	40	Information on notice of abandoned tender for municipal government procurement	Business & industry
18	Tax-related violations - National data set	Credit inquiry	41	Information on correction notice of municipal government procurement	Business & industry
19	Seriously dishonest enterprises in statistics	Credit inquiry	42	Information of signed procurement contracts of district-level governments	Business & industry
20	Information on market access prohibitions	Credit inquiry	43	Information on attachments to municipal government procurement contract	Business & industry
21	Information on administrative punishments	Credit inquiry	44	Information on acceptance of bids for district-level government procurement	Business & industry
22	Illegal fund-raising enterprises	Credit inquiry	45	Information on attachments to district-level government procurement contracts	Business & industry
23	Administrative licenses for enterprises	Credit inquiry			

**Source: Authors' research.**