Right to Data Access in the Digital Era: the Case of China

Yik Chan Chin
Beijing Normal University
Yik-Chan.Chin@bnu.edu.cn
yikchanchin@gmail.com


Abstract

The discussion of epistemic rights is closely linked to the creation and dissemination of knowledge, and data, despite often being treated as a commodity, is actually a form of knowledge. This paper exams the academic debates and Chinese state’s policy about the access to digital data, and demonstrating the lack of epistemic rights manifested in regulating the access to digital data in China and the interplay of global tendencies and local particularities.

The author finds that first, the epistemic right has not drawn attention of Chinese academics, and the closely related concept of right to information is approached from a legal perspective, stressing on the consumer rights to obtain public information and digital platforms’ data rights. Secondly, the right to data access has not been treated as an independent right but as part of the data property right and right to information debates. Thirdly, data is defined as a new factor of production besides land, labor, capital, and entrepreneurship for national economic development in government’s data strategy policy. China’s data access policy has shifted from trading of data ownership right to trading of right to hold data resources, the right to process and use data, and the right to manage data products. To establish a three levels, i.e. national, reginal and industrial levels, data exchange market system would be the next step for academic research and policymaking agenda. Finally, the lack of epistemic rights debate and narrowly defined data has undermined the alternative exploration of public good nature of data, despite the conditional open access to public data, the equality of nonpublic data access might not be included in either academic research or policymaking agenda in the future.
Introduction

China has the second-largest internet market in the world. With the rapid creation and adaptation of digital platforms and e-commerce, the access to, collection and dissemination of data have become the center of academic debate and policymaking. Three factors contributed to these phenomena: 1) the Internet and data are perceived as the important driving force for economic development and an important manifestation of social vitality in China; 2) with the rapid development of the platform economy, the mass production of data has raised governance problems of the storage, transmission, and use of data. 3) the role of digital social media platforms in data access and dissemination has undoubtly strengthened the public demand of government’s acts on the protection of the right to information in China. It is within this context the question of the right to access to data in academic research, policy and regulation becomes the research focus of this chapter.

The primary data used in this chapter include the national Chinese government’s policy and regulations concerning data access, right to information, and data protection. The secondary data include academic literature, research and media reports.

Epistemic Rights and Right to Data Access

According to the definition given by Lani Watson (2021), epistemic rights is closely linked to the creation and dissemination of knowledge - not only about being informed but also about being informed truthfully, understanding the relevance of information, and acting on its basis for the benefit of themselves and society as a whole. In this volume, Hannu Nieminen (Chapter 2) also highlights the equality nature of the epistemic right, such as equality to access to and availability of information and knowledge; and equality in obtaining critical literacy in information and communication.

Data, while often thought of as pure information, is a form of knowledge as it is argued by Gitelman and Jackson (2013) that “raw data is an oxymoron,” and “Data [do] not just exist” (Manovich, 2001). The three concepts of data, information and knowledge are interrelated, but the nature of the relations among them as well as their meanings are debatable. Many scholars claim that data are the raw material for information, and information is the raw material for knowledge (Zins, 2007:479). In this paper, data is defined as set of symbols representing a perception of raw factors. Information is organized data that has been processed into a form that is meaningful to the recipient. Knowledge is understood information (Davis & Olson, 1985; Debons, Horne, & Cronenweth, 1988; Zins, 2007). Digital data is defined as a set of symbols made up of units of binary code that are intended to be stored, processed, and transmitted by digital computers (Zins, 2007:482). Personal data refers to any information which are related to an identified or identifiable natural person (Art. 4 (1), GDPR, 2016). Public data refers to the information collected, produced or paid for by the public or government bodies. Enterprise data refers to data collected and processed by market entities in production and in business activities that do not
involve personal information. Commercial data refers to a type of proprietary data commercialised by a company, and sold by professional data providers with commercial support.

It needs to be imagined as data to exist and function, and the imagination of data involves interpretation. Therefore, data, as a form of knowledge, is created through social processes; its creation and definition therefore involve human agency and interpretation (Berger and Luckmann, 1967: 10; Haggart, 2019). As such, the Chinese academic and policy debates on the access to digital data and its regulation inevitably becomes a social construction process, involving different agencies and interpretations.

Underpinning by the normative criteria of the epistemic right discussed in this volume, this chapter will examine academic debate and national policy of the access to digital data in China. More precisely, this chapter research the conceptualization of right to access to data in China and the related formal and informal rules. It also considers the legitimacy of those rules in relation to the public’s epistemic right to data?

**Right to Access to Data**

In this paper, right to access to data is defined as consisting of two elements: 1) as a right to access to public information which is recognised as an individual human right by many jurisdictions and human rights bodies (Riegner, 2017); 2) as an inclusive right for all members of society to benefit from the availability of data.

Viktor Mayer-Schonberger and Thomas Ramge (2022) define data as a non-rivalrous informational good instead of a physical good, and is a public good for accelerating innovation for the benefit of all. Access to data must align with the fundamental principles of free enterprise and open information flows. They argue that through control of access to data and monopoly of data as raw material, major technology companies could undermine the capacity for innovation as they have less incentive to be disruptive. To address this problem, economic policy must focus on the structural issue of data access and to drastically broaden access to data. Besides, data cannot legally be owned like physical property; affording an exclusive ownership right such as property right to data is impractical due to the difficulties in restricting to a specific purpose or specific users of the use of the data, and trading data in the market is inefficient because the market cannot adequately perform its role as an allocation mechanism. A compulsory opening of the data set is proposed to avoid the concentration of the innovation capacity, to crack down on the information-based domination derived from exclusive access to data. Thus, competitive advantage will rely on extracting insights from data not from access to data. The access mandate includes non-confidential data should be granted for open access, and the direct exchange of data between the data holder and requester is facilitated by an open system of data access.

Purtova (2015) argues that data is not a public good but a rivalrous resource. Without the policy action of assigning property rights including no access and non-disclosure in personal data to the data subject, it will effectively rendering the individual
defenceless in the face of corporate power eroding the autonomy, privacy and right to informational self-determination of the individual.

At the EU level, the EU commission aims a data access for all strategies, that is, data to be available for access to all—whether public or private, big or small, start-up or giant. “Big commercial digital players must accept their responsibility, including by letting Europeans access the data they collect. Europe’s digital transition is not about the profits of the few but the insights and opportunities of the many” (von der Leyen, 2020). The 2022 Data Governance Act allows the creation of common European data spaces for important areas: health, environment, energy, agriculture, mobility, finance, manufacturing, public administration, and skills. Data marketplaces, that is, online platforms where users can buy or sell data - will help new intermediaries be recognised as trustworthy data organisers. Companies, individuals and public organisations can also share personal data for the benefit of society, i.e. data altruism (European Parliament, 2022). Meanwhile, It is suggested that the EU needs to establish a framework for business to government (B2G) data access, and exploring the creation of a cross-EU regulatory framework (European Commission, 2020).

In comparison, in 2022, The World Economic Forum has proposed that Data Marketplace Service Providers (or DMSPs) operate and manage data exchanges: platforms where information, or the right to access certain information under certain conditions, can be traded in an open, efficient and accountable way, participants in data exchanges would trade information collected in a wide range of fields, from healthcare to manufacturing (WEF, 2022).

Academic debate on right to access to digital data in China

The right to access to data has not been treated as an independent right for deliberation but has been part of right to information and data’s property right debates in China.

First, right to access to data is interpreted as part of the personal right to public information if the data is owned by government (Zhang, 2022). There are two theories about its ownership. First, these data should be owned by the public because the source of original data comes from the daily work of government, public financed the collection of data, and the data is ultimately used in people’s daily life, it is public good and its ownership belongs to all of people (R. Huang, Wen and W. Huang, 2018). Second theory propose the data should belong to the state as “the government data ownership is expressed as ownership of individual, the ownership of collective data is rooted in state ownership” (Song and Qiu, 2022).

For nonpublic data, the legal basis of the right to personal information is argued as the right to self-determination of information, any data controller or processor needs to obtain the "expressed consent" of individuals before collecting, obtaining and processing data, and data commercialization that ignores the personal dignity of individuals attached to data should not be accepted. If data protection is not in place, it will damage the rights and interests of individuals and organizations, and even
cause social and economic risks. If overprotected, big data analytics can become impossible (Huang, 2023). But the access right to personal data is not explicitly discussed, and the equality nature of the epistemic right, such as equality to access and availability of information and knowledge, has not drawn much Chinese academic attentions.

Secondly, data access right is treated as part of data’s property right discussion. In other words, contrasting to the EU’s GDPR approach which does not define the ownership of data but regulating the access of data, the Chinese academic debate was concentrating on data’s ownership. This is partly because data is largely not seen as a public good shared by consumers or companies. All activities of data collection, analysis and processes are aimed at unlocking the potential commercial value of data providing personal information and national security are protected (Zhang, 2021). Therefore, to formulate a data trading system supported with data’s ownership right so that data can be traded to generate values became the prioritized pragmatic issue to be addressed. This is also partly triggered by the government’s policy objective on the utilisation of big data, Chinese academic debates are thus heavily policy driven.

Some scholars advocate the establishment of a dual rights structure in which the data subjects own the data and the data processor owns data’s usufruct or operational rights (Shen, 2022; Long, 2017); Or data property rights should be assigned to data companies that collect and process data, and the rights of “sensitive personal data” should be assigned to data subjects (Xu, 2018); Xiaodong Ding (2019) argued against the allocation of the data ownership right to individuals as this would incur extremely high transaction and communication costs and also overtake some of the data rights enjoyed by platforms, making it impossible for platforms to carry out certain normal business activities.

Thirdly, Mei Xiaying (2022), amongst very few of others, supports the public good nature of data and argued that data sharing should be the default position, and control of access to data requires justification because data is a natural public good. The construction of a data control system should be based on the premise of data sharing.

Interestingly, the most recent debate has re-oriented the focus from data ownership to the structural separation of data property right, that is, the property right include three separated rights-data’s holding right, data process and use rights, and data product’s management right. Meanwhile, data sharing is no longer about sharing of original data but sharing of data products. In other words, it is not the original data but the access to data to perform calculation is shared (Huang 2022 & 2023). A researcher at the State Council’s development research center has admitted that the current data trading model is difficult to sustain from both the perspectives of actual needs and government policy (People's Posts and Telecommunications News, 2022). In practice, it is unclear whether and how individuals could have and how to exercise data ownership right, and it is, therefore, impossible to talk about data trading right and data revenue distribution (Zhou et al., 2022). The idea is to use technology such as privacy encryption to separate data ownership from data use right so that data can be used but not shared, data usage can be controlled and measured. And the policy should focus on the development of data services to release data value under the premise of ensuring privacy and security (People's Posts and Telecommunications News, 2022).
Rules governing access to digital data in China

According to incomplete statistics, regulations (drafts) in the name of "data" have sprung up all over the country, and nearly 225 local legislations including 67 local regulations, 158 local departmental rules have been made in China by the end of 2021 (Bai and Li, 2022). The most important element of China’s data strategy policy is that data is officially defined as a new factor of production besides land, labor, capital, and entrepreneurship, and it builds the foundation for the country’s digitalization, connectivity, and AI in the government’s policy documents. To qualify as factor of production, according to a Chinese economist who participated in the government’s data strategy policy drafting, “it must be a must-have basic resources in production of goods and services, data can only qualify as factor of production if it is used in production and business activities and generate significant values” (Huang L., 2023)

First, for the collection and access to the personal data. China’s Personal Information Protection Law (PPL) clearly stipulates that data collector can collect personal information only if it obtains the consent of the individual; or the collection is necessary for the conclusion and performance of a contract; or performance of statutory duties or obligations; or respond to public health emergencies; or conducting news reporting and other acts for the public interest. If collector wants to provide personal information it collected to third parties, it shall inform the individual and obtain their consents. Besides, individual has the right to know, to decide, to rectify, to restrict and refuse the process, to delete, to be forgotten and to obtain an explanation and copy of data.

Also its article 47 established an obligation for data collectors to actively delete personal information if the purpose of process has been achieved, cannot be achieved, or is no longer necessary; or the collector stops providing products or services, or the storage period has expired; or individual withdrawal of consent. As such, on 12th December, 2022, after the State Council announced to seize the use of the health code apps including both the communication travel card and health code, three mobile operators, China Telecom, China Mobile and China Unicom, that are the main data collectors of communication travel card, all announced to delete data related to users synchronously to ensure the security of personal information in accordance with the law. Personal information collected by them after de-identification and anonymization will be provided to relevant government departments in a targeted manner through the joint prevention and control mechanism of the State Council. According to article 4 of the PPL, if the personal information received by government is anonymized, the government agency may independently use such information (Zhang, 2022).

Secondly, data circulation in China is driven by the state’s policies, between 2015 and 2022, the Party, State Council and its ministries have announced a series of policies on the access and trading of data. The policies predominately define data as new factor of production that should be traded according to market mechanism, i.e. to maximize benefits and optimize efficiency based on market rules, prices and competition, to facilitate the country’s economic development.
Table 1. Major Data Policies in China

<table>
<thead>
<tr>
<th>Year</th>
<th>Department</th>
<th>Policy Title</th>
<th>Policy Aims</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>State Council</td>
<td>Action Plan for Big Data Development (促进大数据发展行动纲要)</td>
<td>First national policy document proposed the concept of data trading and provided guidance on data trading market.</td>
</tr>
<tr>
<td>2019</td>
<td>CPC Central Committee</td>
<td>Decision on Several Major Issues Concerning Adhering to and Improving the Socialist System with Chinese Characteristics and Promoting the Modernization of the National Governance System and Governance Capabilities (关于坚持和完善中国特色社会主义制度 推进国家治理体系和治理能力现代化若干重大问题的决定)</td>
<td>Defined data as new factor of production, proposed a mechanism in which the market determines rewards based on contributions.</td>
</tr>
<tr>
<td>2020</td>
<td>CPC Central Committee &amp; State Council</td>
<td>Opinions on Building a Better Market-Allocation System and Mechanism for Factors of Production (中共中央 国务院关于构建更加完善的要素市场化配置体制机制的意见)</td>
<td>Guidance on building a data trading market</td>
</tr>
<tr>
<td>2021</td>
<td>State Council General Office</td>
<td>Overall Plan for Comprehensive Reform Pilot Program of Market-Based Allocation of Factors of Production (要素市场化配置综合改革试点总体方案)</td>
<td>Improving public data sharing mechanism; encouraging enterprises to participate in building trading platforms, and exploring various forms of data trading models.</td>
</tr>
</tbody>
</table>
In 2022, China adopted the most important data policy “Building a Data Base System for Better Use of Data as Factor of Production” to facilitate the compliance and efficient circulation and use of data, to empower economy, to enable sharing of benefits created by digital economy by all people. It is said that the scale of China’s data trading market is nearly trillion RMB, and no one can ignore such an uncapped future market (Fuxi Institution, 2022). The policy sets up of an authorized data access and trading system based on three different types of public, enterprise and personal data. Different access policies are formulated for and applied to each type of data (see Table 2). The property right of data is separated into three rights – i.e. the right to hold data resources, the right to process and use data, and the right to manage data products. Ownership of data is no longer discussed in policy formulation. The government will guide and regulate the data revenue distribution system to reflect both efficiency and fairness (Xinhua News Agency, 2022).

### Table 2. Access Policy on Three Types of Data

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Definition</th>
<th>Access Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Public Data</strong></td>
<td>Data generated by party and government agencies, enterprises and institutions in performing their duties or in providing public services</td>
<td>Strengthen data’s aggregation and sharing, authorized access and management, and interconnectivity; Conditional free access to public data for public interests; Conditional paid access to public data for industrial development; Public data must be provided in the forms of models, products or services but not in original datasets.</td>
</tr>
<tr>
<td><strong>Personal Data</strong></td>
<td>Data bear personal information. Personal information refers to various information related to an identified or identifiable natural person recorded electronically or otherwise, excluding anonymized information.</td>
<td>Data processors can collect, hold, host and use data with valid authorization. Anonymization of personal data is required to ensure information security and personal privacy. Protecting the rights of data subject to</td>
</tr>
</tbody>
</table>
the policy also supports different methods to circulate data, and establish data exchange market systems at national, reginal and industrial sector levels. However, the policy has not addressed how the system can benefit individual data subjects properly. While personal privacy, data security and right to data portability are protected in the policy, how individual data subjects can share the benefits derived from data is not explicitly mentioned and explained.

Conclusions

Access to data as an aspect of epistemic rights has different but similar interpretations in the Chinese and global contexts. First, the epistemic rights in the West academic
literature stresses the sociological nature of the creation and dissemination of information and knowledge. The rights are underpinned by its normative criteria of equal access to and availability of information and knowledge, and use for the benefit of individuals and society as a whole. Therefore, data as a form of knowledge is often defined as non-rivalrous informational good for the benefit of all, open access and sharing of non-confident data is proposed. In the Chinese context, epistemic rights have not drawn attention of Chinese academics, and the closely related concept of the right to information is approached often from a legal perspective, stressing on the consumer rights to obtain public information and digital platforms’ data rights. Data is defined as one kind of factor of production for national economic development.

Interestingly, in China, it is agreed that data has non-rivalrous and non-exhaustive characteristics, the phenomenon of information asymmetry is visible, and data cannot be circulated in the market like land, labor and capital in China, but the public good nature of data has not been considered or recognized in both mainstream academia’s publications as well as in government’s data policies. As a result, the public good and equal access dimensions of data are largely unmentioned in policymaking. Under the premise of protection of national security and personal privacy, data collection, analysis and processes are aiming at unlocking the potential commercial value of data, especially for enterprise data. Therefore, defining the various kinds of property right of data were the focus of academic and policy contestations.

Secondly, like what has been proposed by Viktor Mayer-Schonberger and Thomas Ramge (2022), the recent data access policy in China has shifted from sharing of original data to sharing of data products, from trading of ownership right to trading of holding, process and use, and management rights of data. The establishment of a three-level data trading system at the national, regional and industrial sectors would be the next step for academic research and policymaking, the government will also guide and regulate such development to reflect efficiency of market and fairness of benefit distribution. The public good nature of data and data altruism might not be on either academic research or policymaking agenda, but the open and sharing mechanism of public data are endorsed and encouraged by the government’s policy.

Finally, while the right and interests of data enterprises are the main subject of protection in China’s latest data policy, the power imbalance between the individual and corporations (Purtova, 2015) and the sharing of benefits derived from data with individual users or data subjects have not been addressed.

References


Chongqing Municipal People’s Government. (2021) Interim Measures of Chongqing Municipality on Administration of Government Data Resources. Available at:


Fuxi Institution. (2022) Is the data exchange a good business? Review of Fuxi Institution. Available at:


Huang, L. (2023) The Exploration of Inner Logics of Data as Factor of Production Market. Internal presentation delivered on 4 Jan 2023, Beijing, China.


Ministry of Industry and Information Technology (2016) Big Data Industry Development Plan (2016-2020). Available at:


National People’s Congress. (2020). Article 111 of Civil Code of the People’s Republic of China. Available at:

People’s Posts and Telecommunications News. (2022) ‘What are the data exchange models? What are the important factors that restrict data exchange? Interview with Gao Taishan of the Development Research Center of the State Council’, People’s Posts and Telecommunications News, 30 August. Available at:


Shenzhen City People’s Congress. (2021) Data Regulation of the Shenzhen Special Economic Zone. Available at:


Zhang, X. (2020). ‘Connotation Classification and Ownership Judgement of Data Rights and Interests’, *Shanghai Legal Studies*, 3(01), pp. 338-351.


