



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

**Computer Law
&
Security Review**

Regulating the use of big data: Justifications, perspectives and the Chinese way forward



Wenjie Ding, Li Chen*

Fudan University Law School, Leo Koguan Law Building, 2005 Songhu Road, Shanghai, 200438, China

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Big Data
Non-rivalrous in consumption
Regulatory paradigm shift
Empowerment mode
Behavioural regulation mode

ABSTRACT

Big Data is the collection of data that is non-rivalrous in consumption. Since Big Data is not physically tangible like movable or immovable property and is not subject to restrictive clauses in property law such as the “idea-expression dichotomy” and “fair use and term of protection”, it is fundamentally inappropriate to regard Big Data as an object of real rights. The protection of Big Data requires a paradigm shift from a rights-based “empowerment” approach to that of “behavioural regulation”. Accordingly, the protection of Big Data should not be seen as pitting the approaches of “empowerment” against “behavioural regulation”; rather, it should be seen as achieving coordination between behavioural regulation laws. To effectively balance the interests of data enterprises and the public, this article argues that a competition-based behavioural regulation regime that prohibits unfair competition between data enterprises is more desirable than an intellectual property “rights” regime that protects enterprises’ big data through exclusive rights.

© 2022 Wenjie Ding and Li Chen. Published by Elsevier Ltd. All rights reserved.

Introduction

On 11 March 2021, *The Outline of the Fourteenth Five-Year Plan for National Economic and Social Development of the People's Republic of China and Vision 2035* was approved at the 4th session of the 13th National People's Congress with the objectives of “accelerating digital development and building a Digital China” and “establishing sound market rules for data factors”. While China currently lacks a specific protection mechanism for Big Data (with the development of legislation seemingly envisioned under this Plan), existing laws can, to some extent, fill the gap caused by the legislative delay. Firstly, although data collection is somewhat automatic and mechan-

ical, the methods by which data is collected and processed may be regarded as an original compilation work in its entirety and may therefore be granted copyright protection.¹ Secondly, Big Data can be effectively protected as a trade secret subject to the requirements of secrecy, commercial value, confidentiality measures, and commercial information.² Therefore, the lacuna in Chinese law regarding the protection of Big Data mainly refers to data collection that is made public and by methods that are not original.

¹ See *Jinan White Rabbit Information Co Ltd v. Foshan Dinrong Software Technology Co Ltd Copyright Dispute*, Foshan Intermediate People's Court, Guangdong Province (2016) Yue 06 Final Civil Judgment No. 9055.

² See *Quzhou Wanlian Network Technology Co., Ltd. v. Zhou Huimin et al. Infringement of Trade Secrets Dispute*, Shanghai High People's Court (2011) Shanghai High People's Third (Zhi) Final Civil Judgment No. 100.

* Corresponding author.

E-mail addresses: dingwenjie@fudan.edu.cn (W. Ding), lichen@wustl.edu (L. Chen).

Unfair competition caused by the misappropriation of Big Data (such as user reviews,³ Weibo users,⁴ real-time traffic,⁵ and other derivatives)⁶ has led to demands for regulation.⁷ As a stopgap measure, the Chinese courts normally address this by invoking the general provisions in Article 2 of the *Chinese Anti-Unfair Competition Law*. Yet, this legislation only contains highly abstract and overarching general provisions which do not concretely articulate what the main approach to Big Data protection should be.⁸ In this regard, academic literature generally favours the enactment of special legislation to govern Big Data. However, the prevailing research paradigm is still confined to the traditional approach of “empowerment,” as reflected in two aspects. First, in the context of mainstream civil law scholarship, scholars have predominantly sought to protect Big Data with a real rights-based regime, neglecting the non-rivalrous nature of Big Data consumption.⁹ Similarly, from the intellectual property perspective, some scholars have likewise attempted to devise rules for Big Data protection within the framework of intellectual property law and its associated rights.¹⁰

The problem with the empowerment approach is its overemphasis on data enterprises’ entitlements to Big Data, which overlooks the dynamic equilibrium between data protection, data sharing, and data security. This is especially true with the emergence of monopolistic practices such as “picking one from two” (where a data platform requires users to exclusively use its platform), big data swindling by data enterprises, as well as data security risks arising from offshore listings of data enterprises. Therefore, it is increasingly challenging to regulate data sharing and its safe use.¹¹ Indeed, this

is why state administrative bodies are gradually tightening up their supervision of data companies in recent years. For instance, in April and October 2021, the State Administration for Market Regulation handed down administrative penalties against Alibaba and Meituan for abusing their dominant market position by engaging in “picking one from two” activities¹² because it “imped[ed] the free flow of market resources”. Another example is the cybersecurity review conducted by the Cyberspace Administration of China in July 2021 against Didi Chuxing for possibly disclosing its users’ personal information, travel data, and road information to foreign regulators during its listing on the New York Stock Exchange.¹³ The behaviour of data enterprises undoubtedly serves as a warning about the perils of the empowerment approach.

Indeed, the potential drawbacks inherent in the empowerment approach prompt us to ponder more deeply: what underpins the protection of Big Data? Does the legal protection of Big Data unduly restrict its circulation? What is the best solution for the protection of Big Data? To that end, this article aims to clarify the rationales underpinning the protection of Big Data, point out the lacuna of the dominant real rights-based empowerment approach, and elaborate on the significance of the paradigm shift from empowerment to behavioural regulation in guiding the development of Big Data legislation in China.

The fundamentals of big data

Big data is non-rivalrous in consumption

From the viewpoint of institutional economics, Big Data is the collection of data that is non-rivalrous in consumption. This means that an increase in the number of persons consuming such data does not require an increase in the number of goods provided. In other words, the marginal cost of additional consumers is near zero¹⁴ and the consumption of Big Data by one firm normally does not prevent others from consuming the same. Consequently, Big Data can be consumed by countless firms and create diverse value: the same Big Data can be exploited for multiple purposes, resulting in different value-added services or derivative applications. Such varied uses can also lead to more complex and advanced applications.¹⁵

practice that triggered an official probe? 24 Dec 2020, <https://www.scmp.com/tech/enterprises/article/3115300/alibaba-antitrust-investigation-what-picking-one-two-practice>. China Signals Broad Clampdown on Company Data, Offshore Listings, <https://time.com/6078125/china-crackdown-didi>.

¹² See the Decision on Administrative Penalties by State Administration for Market Regulation (SAMR) (2021) No. 28; and the Decision on Administrative Penalties by SAMR (2021) No. 74.

¹³ Keith Zhai, Frances Yoon, Beijing Blocks Merger, Tightens Data Rules as Post-Didi Crackdown Speeds Up, WSJ, 10 July 2021, <https://www.wsj.com/articles/beijing-blocks-merger-tightens-data-rules-as-post-didi-crackdown-speeds-up-11625898515>.

¹⁴ See Ji Hailong, “The Private Law Positioning and Protection of Data”, *Chinese Journal of Law*, No. 6, 2018, p. 81.

¹⁵ See Wu Weiguang, “A Critique of the Theory of Private Rights Protection of Personal Data Information in Big Data Technology”, *Political Science and Law*, No. 7, 2016, p. 118.

³ See *Shanghai Hantao Information Consulting Co., Ltd. v. Beijing Baidu.com Technology Co., Ltd. Unfair Competition Dispute*, Shanghai Intellectual Property Court (2016) Jing 73 Final Civil Judgment No. 588.

⁴ See *Beijing Weimeng Chuangke Network Technology Co., Ltd. v. Beijing Taoyou World Technology Co., Ltd. Unfair Competition Dispute*, Beijing Intellectual Property Court (2016) Hu 73 Final Civil Judgment No. 242.

⁵ See *Shenzhen Gummi Technology Co., Ltd. v. Wuhan Yuan-guang Technology Co., Ltd. Unfair Competition Dispute*, Shenzhen Intermediate People’s Court, Guangdong Province (2017) Yue 03 Civil Judgment No. 822.

⁶ See *Taobao (China) Software Co Ltd v. Anhui Vision Information Technology Co Ltd Unfair Competition Dispute*, Hangzhou Intermediate People’s Court, Zhejiang Province (2018) Zhe 01 Final Civil Judgment No. 684.

⁷ China Seeks Better Cross-border Control of Big Data with New Plan, available at <https://www.reuters.com/technology/china-urges-stronger-cross-border-security-big-data-2021-11-30>.

⁸ See Jiang Ge, “Relationship Between Intellectual Property Laws and the General Clause in Countering Unfair Competition Law”, *Chinese Journal of Law*, No. 2, 2019, p. 118.

⁹ See Long Weiqiu, “On the Construction of New Data Property and its System Structure”, *Tribune of Political Science and Law*, No. 4, 2017, pp. 63-77; Cheng Xiao, “Personal Data Rights in the Era of Big Data”, *Social Sciences in China*, No. 3, 2018, pp. 102-122; Shen Weixing, “On Data Usufruct”, *Social Sciences in China*, No. 11, 2020, pp. 110-131.

¹⁰ See Cui Guobin, “Towards a Theory of Limited Exclusive Right to Big Data”, *Chinese Journal of Law*, No. 5, 2019, pp. 3-24.

¹¹ Jane Zhang and Celia Chen, *Alibaba antitrust investigation: what is the ‘picking one from the two’*

However, some Chinese scholars fail to recognize this fundamental characteristic of Big Data. Instead, they assert that Big Data has the required independence and nature to be regarded as an object of property rights under civil law.¹⁶ They compare Big Data to electricity, heat, sound, light and space, all of which are novel objects of property rights in modern civil law. For example, a well-established Chinese scholar, in exploring the object of data property rights, argued that “data existing in physical forms of light, electricity, magnetism, etc., although invisible to the naked eye, is an objective being and its objective existence suffices to make it an object of real rights.”¹⁷ Modern property law considers the above-mentioned physically intangible things as objects of real rights¹⁸ because they all share the common characteristic of being rivalrous in consumption. It is regarded as an extension of tangible property and classified as *res corporales*.¹⁹ Therefore, unlawful seizure and theft of electricity can also constitute misappropriation and larceny.²⁰

Theoretical conundrum of big data protection

Real property rights are justifiable in respect of tangible objects, whether movable or immovable because tangible things are rivalrous in consumption. They cannot be used by others once it comes into the possession of one person. For example, if A possesses an orange, no one else can consume the same orange. As a result of this characteristic, it is imperative that persons are able to hold exclusive property rights over a tangible thing; this is for two reasons. First, without such rights, one has to constantly fear that their hard work might be stolen by others.²¹ Secondly, by granting private property rights over a common resource, there is an incentive for its owner to avoid over-exploiting the resource and to ensure its availability to future generations.²² Otherwise, this is likely to lead to the “tragedy of the commons” problem,²³ which posits that people are likely to exhaust common resources in pursuit of their own interests in the absence of real property rights.

In contrast, this rationale for according real property rights does not apply to Big Data which is non-rivalrous in consumption. Unlike tangible objects such as moveable or immovable property, Big Data does not suffer from the problem of scarcity. Even unauthorised access of Big Data by a third party would not impede the enterprise’s own use. The misappropriation

of Big Data by others does not deprive the data enterprise of its freedom to reproduce and sell the data, although it might potentially lead to a loss in profits. Thus, it is clear that this fundamental difference between Big Data and traditional tangible objects is that Big Data can never be over-consumed to the point of depletion and that the “tragedy of the commons” problem in the realm of real rights cannot by itself justify the protection of Big Data.²⁴

Critique of the real rights-based empowerment mode

The internal logic of the real rights-based empowerment mode

The majority of Chinese civil law scholars resort to a property rights-based empowerment approach to protect Big Data. Although civil law scholars disagree on the origin of Big Data rights, they share a common understanding: Big Data protection is justified by the dual theory of “data ownership” and “lawful data collection”. However, given the substantial difference between Big Data and traditional tangible objects, this segment will dive deeper into the labour theory of data ownership and the principle of lawful data collection to determine whether these principles can be directly transplanted into the theory of data property rights.

The jurisprudential flaw in the labour theory of data ownership

The labour theory of data ownership is derived from the labour theory of property in legal philosophy; the latter was formulated by the English philosopher John Locke.²⁵ However, this theory is not fully applicable to Big Data.

Although Locke’s *Two Treatises of Government* only deals with property rights in a general sense, his labour theory of property furnishes civil law scholars with the intellectual foundation for a real rights-based empowerment approach, most notable of which is the theory of value. Based on the theory of value, because a data enterprise invests enormous capital, material and human resources in collecting, processing and analysing data, it should be rewarded with and be entitled to data property rights.²⁶ While somewhat persuasive, the theory of value cannot be directly transplanted to the field of data property rights for three reasons.

First, the reward for the efforts of the data enterprise need not take the form of a real right.²⁷ For example, a proportionate amount of social bonus or public funding could qualify as a reward. The reward could also be social or moral in nature – such as honour, acknowledgement, status or power.

¹⁶ See *supra* note 9, Cheng Xiao, p. 107.

¹⁷ *Supra* note 9, Shen Weixing, p. 122.

¹⁸ See Wang Liming, “Discussing Some Issues of Property Law Legislation”, *Tribune of Political Science and Law*, No. 4, 2001, p. 11.

¹⁹ See Qu Maohui, “Two Foundational Questions on the Object of Property Rights”, *Present Day Law Science*, No.2, 2005, p. 19.

²⁰ See Wang Liming, *A Treatise on the Property Law (Revised)*, China University of Political Science and Law Press, 2003 edition, p. 32.

²¹ See Harold Demsetz, “Toward a Theory of Property Rights”, *American Economics Review*, Vol.57, 1967, p. 347.

²² See 山根崇邦「知的財産権の正当化根拠論の現代的意義(1)」知的財産法政策学研究会28号(2010年)214頁。(See Takakuni Yamane, “Contemporary Significance of the Debate on the Rationale for Justifying the Intellectual Property Rights (1)”, *Intellectual Property Law and Policy Journal*, No. 28, 2010, p. 214.)

²³ Garrett Hardin, “The Tragedy of the Commons”, *Science*, Vol.161, 1968, p. 1243.

²⁴ See Ding Xiaodong, “On Legal Protection of Enterprises’ Data: Analysis Based on the Legal Nature of Data”, *Science of Law (Journal of Northwest University of Political Science and Law)*, No. 2, 2020, p. 96.

²⁵ See Locke, *The Second Treatise of Government*, translated by Ye Qifang and Zhai Junong, The Commercial Press, 1964, pp. 18-27.

²⁶ See *Supra* note 9, Cheng Xiao, p. 117, Shen Weixing, p. 128.

²⁷ See Edwin C. Hettinger, “Justifying Intellectual Property”, *Philosophy & Public Affairs*, Vol.18, 1989, pp. 41-42.

Second, granting data property rights to enterprises as remuneration for their labour might be inappropriate because it might not be commensurate with the work done by the data enterprise. The labour and time involved in data creation may not be proportionate to the value of the results, as circumstances beyond human control often influence the latter.

Third, the justifications for according property rights under Locke's labour theory of property is inapplicable to Big Data. Locke's theory is based on the inviolability of self-ownership — an exclusive right of people to their bodies, faculties, talents, and energies,²⁸ which restricts the physical use of the object by others. Such a restriction makes sense for movable or immovable property, which is rivalrous in consumption, and is crucial to the efficient exercise of the rights to use, benefit and dispose of ownership: a price necessary for social coexistence.²⁹ However, the non-rivalrous nature of Big Data consumption, which allows Big Data to be used by more persons for more purposes, means that the cost to society to limit the use of Big Data is far greater than what it would be for objects that are rivalrous in consumption.

Misconceptions about the principle of lawful data collection

The dominant view in civil law treats the principle of lawful data collection as another legitimate basis for data property rights.³⁰ However, this principle merely exempts data enterprises from infringement liability and is not a justification for the entitlement of data enterprises to data that is collected. Specifically, where data is collected lawfully, i.e., with the informed consent of the data subject, the data enterprise is not liable in tort. Conversely, if data is collected unlawfully – i.e., where it infringes on the right to privacy or the right to reputation of the subject, the data enterprise is liable for the infringement, but this does not preclude the data enterprise from prohibiting the unlicensed use of such data by others.

Another issue with the principle of lawful data collection is that it is premised on the real rights-based empowerment approach. This is evidenced by the rights given to data enterprises – to possess, use, benefit from and dispose of data collections or other data products lawfully collected.³¹ This is a problem because Big Data is non-rivalrous in consumption and does not share the same attributes as objects of real rights. For the owner of an object of a real right to exercise dominion and complete use of it, the owner must be entitled to exclude dominion or use by others. For example, the owner cannot ride a bicycle if someone has stolen it, which justifies the right to claim the return of a corporeal object as an effect of ownership. In contrast, for Big Data, a third party may exploit it anywhere and anytime without needing to seize and

possess it from the data enterprise. Meanwhile, the unauthorised use by the third party is no barrier to the enterprise's own use of big data. Therefore, to realise the legal effect of restitution, a negative exclusive right to prohibit use by others will suffice.

The tug of war between the protection and free use of data

The real rights-based empowerment approach lies on the most extreme end of the spectrum amongst the methods of employing property rights for big data protection. This approach exclusively allocates the benefits of Big Data to data companies, resulting in undue restrictions on the transmission and use of data in the public domain.

Existing copyright laws can somewhat ameliorate the problem by safeguarding the copyright of data enterprises while avoiding undue restrictions on the freedom of data access in the public domain. In particular, the balancing mechanisms in Chinese copyright law, such as the “idea-expression dichotomy” and “fair use and term of protection,” can strike a delicate balance between these two goals. First, the idea-expression dichotomy limits copyright protection to those parts of a database which are original expressions; conversely, individual data, simple facts, and abstract ideas devoid of originality fall into the public domain.³² Second, there are limited exceptions to copyright protection, including fair use provisions that allow for private copying and use of data provided that it is appropriately quoted. Third, the duration of copyright is limited; after a certain time period, databases with originality will be accessible in the public domain. On balance, even though these balancing mechanisms do not eradicate the tension between the protection and free use of data,³³ they go some way to promote the free transmission and use of data.

In contrast, property law cannot strike the right balance between the protection and use of data. As alluded to earlier, property law traditionally regulated the physical use of tangible things that are rivalrous in consumption. This approach, however, does not apply to Big Data due to its intangible and non-rivalrous nature. If strong property rights were to be granted over Big Data, it would impose unreasonable restrictions on the freedom of data utilisation because the potential use of Big Data is far greater than tangible things that are limited by their physical existence in space. The restrictions of property law are all the more unreasonable given that it does not carve out limitations or exceptions, unlike copyright law which has balancing mechanisms such as the “idea-expression dichotomy” and “fair use and terms of protection”.

²⁸ See 森村進「財産権の理論」(有斐閣, 1995年)18頁.(See Susumu Morimura, *Theory of Property Rights*, Yuhikaku, 1995, p. 18.)

²⁹ See 山根崇邦「知的財産権の正当化根拠論の現代的意義(3)」知的財産法政策学研究31号(2010年)127頁.(See Takakuni Yamane, "Contemporary Significance of the Debate on the Rationale for Justifying the Intellectual Property Rights (3)", *Intellectual Property Law and Policy Journal*, Vol. 31, 2010, p. 127.)

³⁰ See *Supra* note 9, Cheng Xiao, p. 118, Shen Weixing, p. 128.

³¹ See Yao Jia, "Guidelines for the Utilization of Corporate Data", *Tsinghua Law Review*, No. 3, 2019, p. 123.

³² See Melville B. Nimmer, "Does Copyright Abridge the First Amendment Guarantees of Free Speech and Press?", *U. C. L. A. Rev.*, Vol. 17, 1970, pp. 1189-1193.

³³ See 比良友佳理「デジタル時代における著作権と表現の自由の衝突に関する制度論的研究(4)」知的財産法政策学研究48号(2016年)61-95頁.(See Yukari Hira, "Copyright and Freedom of Expression in the Digital Era (4)", *Intellectual Property Law and Policy Journal*, Vol. 48, 2016, pp. 61-95.)

Justifications of the behavioural regulation mode

The theoretical basis of the behavioural regulation mode

As a collection of data that is non-rivalrous in consumption, Big Data fits into what intellectual property rights (“IPRs”) seek to regulate. Accordingly, intellectual property law scholars have attempted to lay down broad rules for the protection of enterprise big data within the framework of intellectual property law. There are two differing explanations for this. The prevailing view is based on the notion of an intellectual property “rights” regime that awards exclusive rights that are less absolute than real property rights, to Big Data enterprises.³⁴ The other explanation sees intellectual property protection of Big Data as a “Behavioural Regulation” regime that governs the benefits accorded to data enterprises through behavioural controls.³⁵ However, it is overly simplistic to regard the two explanations as being dichotomous or even in competition with each other.

Examining the difference between “rights” and “behavioural regulation”

Scholars generally believe that intellectual property law in a narrow sense refers to laws protecting “rights” such as copyright, patents, and trademarks. Yet, the essence of IPRs is “behavioural regulation” and not “rights” over intangible “objects” such as works, inventions, and Big Data. This approximates competition law, which is also a form of behavioural regulation concerning “acts” such as counterfeiting and confusion, trade secret infringement, and unfair competition on the internet.

The notion of “intellectual property” is sometimes misunderstood due to the limitations of languages or metaphors employed in describing it.³⁶ Many mistakenly understand “intellectual property” to denote the right to intangible “objects”. However, this presupposes that certain intangible “objects” in the world could be harnessed by human beings³⁷ and that there is a clear distinction between intangible “objects” and the “act” of using them, which is not the case. Intangible “objects” are concepts in people’s minds,³⁸ and the distinction between intangible “objects” and the “act” of using them is often a product of characterisation. As Yoshiyuki Tamura proposed, the current copyright law regime characterises the act of dis-

tributing specific computer software via the internet as a distribution (act) of the software work (intangible object) to the public. However, in the current patent law regime, the nature of acts should be determined according to the precise patent claim made. If a specific computer software were to be treated as an invention, the online dissemination would be regarded as the implementation “act” of the invention; whereas, if the method of online dissemination was regarded as the invention itself, then the method would be regarded as the invention itself and it would be regarded as the relevant intangible “object”.³⁹ This shows that the notion of intangible “objects” is a way to label diverse human behaviours into concepts such as “works”, “inventions”, and “Big Data”.⁴⁰

Reconceptualizing the legitimate basis of big data protection

Despite some controversy, the essence of IPRs is “behavioural regulation” – the use of legal means to regulate data use; it is not focused on “rights” over intangible “objects”. This means that the mere fact that Big Data has property value cannot itself justify limiting the freedom of data use in the public domain or automatically lead to the conclusion that Big Data deserves legal protection. The legal protection of Big Data requires twin bases to justify the regulation of data use by data enterprises: namely, a utilitarianism-based incentive theory as a positive basis and a natural rights theory as a negative basis.

First, the positive basis for Big Data protection is utilitarianism. Enterprises invest hefty sums of financial, material and human resources in collecting, processing, and analysing data. Should such data be left to competitors to free-ride upon, it will inevitably dampen the enthusiasm of enterprises in developing products and exploring the potential value of Big Data, eventually hampering the development of the digital economy and Big Data industry.⁴¹ Thus, the positive basis for Big Data protection concerns not only the interests of the enterprises themselves but also those of a broader majority. Put differently, failing to regulate certain kinds of free-riding acts will drastically reduce the number of enterprises dedicated to Big Data exploitation, which ultimately harms the public interest. Since public welfare is at stake, utilitarianism can act as a positive basis for limiting the free use of data.

Second, the natural rights theory adds a negative basis for Big Data protection. While utilitarianism takes public welfare as a positive basis, the ends-means mentality of regulating free use of data for the greater good in and of itself still falls short as a conclusive basis of legitimising Big Data protec-

³⁴ See Lv Bingbin, “On the Right of Network Users to the ‘Data’—and on Industrial Policy and Balancing of Interests in the Cyber Law”, *Science of Law (Journal of Northwest University of Political Science and Law)*, No. 6, 2018, p. 62.

³⁵ See Ye Jinqiang, “The Gain and Loss of the ‘Civil Rights’ Chapter of General Provisions of Civil Law”, *Peking University Law Journal*, No. 3, 2017, p. 650.

³⁶ See 松浦好治『法と比喩』(弘文堂, 1992年)3-4頁。(See Yoshiharu Matsuura, *Law and Metaphor*, Koubundou, 1992, pp. 3-4.)

³⁷ See Ma Junju and Mei Xiaying, “On the Theoretic and Legislative issue of Intangible Property”, *Chinese Legal Science*, No. 2, 2001, p. 104; Xiaolan Yu, Yun Zhao, Dualism in data protection: Balancing the right to personal data and the data property right, *Computer Law & Security Review*, Volume 35, Issue 5, 2019.

³⁸ See Peter Drahos, *A Philosophy of Intellectual Property*, Dartmouth Publishing Company Limited, 1996, p. 18.

³⁹ See Yoshiyuki Tamura, “Conceptual Fallacies Behind the Idea of an Area Without Protection of Intellectual Works”, translated by Li Yang and Xu Qing, *Jurist*, No. 4, 2010, p. 119.

⁴⁰ Wendy J. Gordon (田辺英幸訳)「INTELLECTUAL PROPERTY」知的財産法政策学研究11号(2006年)6-7頁。(See Wendy J. Gordon, “Intellectual Property”, translated by Toshihide Watabe, *Intellectual Property Law and Policy Journal*, No. 11, 2006, pp. 6-7.)

⁴¹ See Li Yang and Li Xiaoyu, “The Definition of the Nature of Enterprise Data Rights and the Construction of Its Protection Model in the Age of Big Data”, *Academia Bimestrie*, No. 4, 2019, p. 181.

tion.⁴² This is where the natural rights theory bolsters the argument for Big Data protection; it posits that enterprises create economic value during the collection, processing, and analysis of raw data, which militates toward Big Data protection. While this negative argument cannot justify Big Data protection in and of itself, it supplements the more persuasive positive basis of Big Data protection, i.e. utilitarianism.

Paradigm shift: from empowerment to behavioural regulation

As argued earlier, the essence of big data protection is to regulate the freedom to use data in a manner consistent with the public good. Therefore, Big Data protection should not focus on the paradigm of pitting “empowerment” against “behavioural regulation”. Rather, discussions of Big Data protection should focus on a more detailed analysis from the perspective of behavioural regulation.

Perceived advantages of the behavioural regulation approach

Analysing Big Data protection from the perspective of behavioural regulation is preferable for three reasons.

First, the essence of Big Data protection is the use of law to limit the non-rivalrous data use of a public good, which precisely describes what behavioural regulation seeks to do. Absent legal protection, Big Data can be freely used by anyone as a public good, which might disincentivise data enterprises from collecting data.

Second, behaviour regulation balances the need to incentivise data enterprises to continue their work with interest in allowing the public to also utilise the data that is collected. Each interest would have to be weighed against another. The mere fact that Big Data qualifies as a “legitimate right” does not in itself mean that it deserves legal protection. Conversely, the singular fact that a competitor can free-ride by unauthorized data crawling does not *ipso facto* necessitate legal regulation of the competitive behaviour. Society at large benefits from such free-rider behaviour: For example, including keywords of other undertakings’ products in the web page title and content introduction, as long as the expression does not harm the interests of consumers and does not substantially hinder other undertakings from obtaining trading opportunities, it will not constitute unfair competition.⁴³ In weighing the competing interests, an instance where regulation is needed from a utilitarian perspective is when free-riders harm the “legitimate rights” of Big Data developers and erode the incentives for Big Data development.⁴⁴

Third, protecting Big Data from the perspective of behaviour regulation avoids the pitfalls of a real-rights empowerment approach that grants ownership rights over Big Data.

Under the empowerment approach, the original data holders should also have avenues for recovery of proceeds if persons other than the original data holders were to use the data. However, such a legal regime in which the majority has a right to claim over Big Data is likely to incur very negative results. In this regard, to facilitate third parties’ obtaining licences for data use, it would be more advisable to grant data enterprises a negative exclusive right to the data.

Division of competencies amongst behavioural regulation laws

Since the essence of Big Data protection is behavioural regulation and not rights over intangible “objects”, it should be seen as requiring coordination between intellectual property laws, anti-unfair competition laws, and other behaviour regulation laws.

As mentioned above,⁴⁵ the current *Chinese Copyright Law* prevents undue restrictions on the freedom to use original data collections (including databases) through restrictive provisions such as the “idea-expression dichotomy” and “fair use and term of protection”. However, an open question is whether non-original Big Data made public can still be protected by establishing rules under the framework of *Chinese Copyright Law*.

There is unlikely to be a clear-cut answer to this question. An example can be gleaned from South Korea, where neighbouring rights protection was explicitly granted to non-original databases in the 2003 legislative revisions to the *Korean Copyright Act*.⁴⁶ This was, however, a controversial development that was strongly criticised by Korean scholars; they pointed out that it is inconsistent with the fundamental purpose of copyright law and unconstitutional for data enterprises to enjoy exclusive rights like copyright over a non-original database.⁴⁷

Therefore, legal protection of non-original Big Data requires an institutional design that is compatible with the intellectual property system. After all, the role of copyright law is to recognise the protection of intellectual output that is expressed in an original way. In contrast, Big Data is valued primarily for its capacity to generate information of high economic value by rapidly collecting and analysing large amounts of data,⁴⁸ which itself often requires significant investment. There is thus an apparent mismatch between the practical value of Big Data and the notion that copyright law mainly protects original expression.

Another complication is that providing copyright protection for non-original Big Data discourages the reasonable circulation and use of data, and this goes against the interest in

⁴⁵ *Ibid*, Part III.2.(ii)

⁴⁶ For the background and process of the relevant legislation, see: , , ,2007,902-905 . (See Jong Sang Jo, *Commentary of Copyright Law*, Pakyoungsa, 2007, pp. 902-905.)

⁴⁷ , , , ,2011,310 . (See Jong Sang Jo and Park Jun Seok, *Intellectual Property Law*, Hongmoonsa, 2011, p. 310.)

⁴⁸ See 田村善之「ビッグ データの不正利用行為規制の新設について—平成30年不正競争防止法改正の解説—」法学教室462号(2019年)65頁.(See Yoshiyuki Tamura, “Big Data Unauthorized Use Regulations: Explanation of the 2018 revision of the Unfair Competition Prevention Law”, *Hogaku Kyoshitsu*, No. 462, 2019, p. 65.)

⁴² See Jeremy Waldron, “From Authors to Copiers: Individual Rights and Social Values in Intellectual Property”, *Chi-Kent L. Rev.*, Vol. 68, 1993, pp. 841, 862-86.

⁴³ See Zhou Xiping, “The Application of the ‘Free-Rider’ Theory in the Protection of Commercial Labels—From the Case of Unfair Competition of Keywords”, *Legal Science*, No.5, 2017, p. 138.

⁴⁴ See 田村善之「機能的知的財産法の理論」(信山社, 1996年)14頁.(See Yoshiyuki Tamura, *Theory of Functional Intellectual Property Law*, Shinzansha, 1996, p. 14.)

promoting freedom of action. It is more desirable to confine the scope of protection of non-original Big Data in a manner that still incentivises data enterprises to behave competitively and for a period of time sufficient for those enterprises to recover their invested capital. A competition law approach for Big Data protection would create a more effective incentive for data enterprises and ensure the free flow and use of data.

The proposed Chinese approach to the behavioural regulation mode

Lessons from the past: EU *sui generis* right and data producer's right

The EU has adopted the intellectual property "rights" approach toward regulating Big Data. The *Directive on the Legal Protection of Databases* issued by the European Union in March 1996 ("the EU Database Directive") stipulates a "double protection" system for databases. First, copyright protection is granted over databases that constitute the author's intellectual creation by reason of the selection or arrangement of their contents. Second, a *sui generis* right is granted over databases that constitute a substantial investment (qualitatively or quantitatively), regardless of their originality.⁴⁹

However, the availability of a *sui generis* right over databases is problematic. The European Commission issued the first evaluation report in 2005⁵⁰ to assess whether the Directive had achieved its policy objectives and whether the newly created *sui generis* right had impeded fair competition.⁵¹ This was after all the EU countries had implemented the EU Database Directive in 2001 and the European Court of Justice had issued its first judgement on the *sui generis* right in 2004.⁵² In the report, the European Commission was critical of the *sui generis* right; the Commission highlighted problems such as the unclear scope of the right, the perverse incentive to lock up information, as well as the failure to achieve any measurable impact on data production and the development of the database industry. It compared the database industry in the EU with that of the United States, where no such *sui generis* rights are provided and where the database industry continues to prosper. Therefore, the *sui generis* right has not served its intended purpose and requires revision.⁵³

⁴⁹ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases, Art.7, 9, 10. See Jörg Hladjk, "The Protection of Databases under US and EU Law – Sui Generis Right as an Appropriate Concept? – Part I – US Law", *Computer Law & Security Review*, Vol. 20, 2004, pp. 288-292. Gillian Bull, "How Will the EU Database Directive and the UK Regulations Impact on Database Use? – Part II", *Computer Law & Security Review*, Vol. 14, 1998, pp. 367-376.

⁵⁰ DG Internal Market and Services Working Paper, First Evaluation of Directive 96/9/EC on the Legal Protection of Databases, on 12 December 2005.

⁵¹ Evaluation of Directive 96/9, 1.1.

⁵² *The British Horseracing Board Ltd v. William Hill Organization Ltd*, C-203/02(2005)1 C.M.L.R. 15(UK); *Fixtures Marketing Ltd. v. Svenska Spel AB*, C-338/02(Sweden); *Fixtures Marketing Ltd. v. Oy Veikkaus Ab*, C-46/02(Finland); *Fixtures Marketing Ltd. v. Orignismos Prognostikon Agonon Podosfairou AE*, C-444/02(Greece).

⁵³ Evaluation of Directive 96/9, 5.

The debate on Big Data protection has not ceased. A *Digital Single Market Strategy for Europe*, published by the European Commission in May 2015,⁵⁴ states in no uncertain terms that the aforementioned issues would be one of its future policy topics, and relevant discussions have intensified accordingly. In addition, in January 2017, the European Commission issued the newsletter titled *Building a European Data Economy*, in which it considered granting rights to data producers (i.e., the owners or long-term users/lessees of data-generating devices) over non-personal data, such as a right to use and authorise the use of non-personal data.⁵⁵ This generated a lot of dissent amongst intellectual property law academics in the EU who argued against granting rights to data producers. For example, the Max Planck Institute for Innovation and Competition in Germany has proffered a very pointed view, stating that the creation of exclusive rights on data may hinder data circulation and result in data monopoly.⁵⁶

A valuable attempt: Japan regulates unfair competition in shared data of limited access

Japan is the first country to protect Big Data through anti-unfair competition law, and it has been working on preparations for Big Data legislation since 2017. Its basic policy is four-fold: to balance the interests of data providers and users; promote the overall circulation and utilization of data; focus on violations and data transactions between civil subjects; and only regulate those egregious behaviours to the minimum extent necessary.⁵⁷ Under the guidance of this policy, considering that an intellectual property "rights" mode similar to *sui generis* rights of the EU may impede data circulation, Japan believes that it is more appropriate to protect competitive data interests through anti-unfair competition law.⁵⁸ Therefore, in May 2018, the clause which provides for shared data of lim-

⁵⁴ Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee & The Committee of The Regions, *A Digital Single Market Strategy for Europe*, online: EUR-Lex < <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52015DC0192>>.

⁵⁵ See Václav Janeček, "Ownership of personal data in the Internet of Things", *Computer Law & Security Review*, Vol. 34, 2018, pp. 1039-1052.

⁵⁶ Josef Drexler, Reto M. Hilty et al., *Data Ownership and Access to Data: Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate*, Max Planck Institute for Innovation & Competition Research Paper No.16-10.

⁵⁷ See 産業構造審議会知的財産分科会不正競争防止小委員会「データ利活用促進に向けた検討中間報告」(経済産業省, 2018年).(See Ministry of Economy, Trade and Industry, "Practical response to the 2018 Revised Anti-Unfair Competition Law", 2018, <https://www.meti.go.jp/report/whitepaper/data/20180124001.html>.)

⁵⁸ See 知的財産戦略本部「新たな情報財検討委員会報告書—データ・人口知能(AI)の利活用促進による産業競争力強化の基盤となる知財システムの構築に向けて—」(2017年)2頁.(See Intellectual Property Strategy Headquarters, "New Information Goods Review Committee Report", 2018, p. 2, https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kensho_hyoka_kikaku/2017/johozai/houkokusho.pdf.)

ited access was incorporated into Japan's *Law Against Unfair Competition*⁵⁹ to regulate the improper use of Big Data.

There are two facets of this law that are worthy of note. First, the target of protection under Japan's Big Data provisions is shared data of limited access. Shared data of limited access is technical or business information accumulated in significant quantities and managed electromagnetically (except as trade secrets) by electromagnetic methods (electronic, magnetic and other methods not perceivable by humans) and provided to a specific person for profit.⁶⁰ Second, with respect to the types of unfair competition regulated, Article 2(1) of the law provides that the improper acquisition, use or disclosure ;⁶¹ use or disclosure in significant breach of good faith ;⁶² and improper acquisition, use or disclosure by a transferor⁶³ of shared data of limited access constitute unfair competition. The right to stop the infringement⁶⁴ and to claim damages are provided for the aforementioned behaviours.⁶⁵

The Japanese approach is a useful attempt to regulate the improper acquisition, use or disclosure of shared data of limited access through anti-unfair competition law.⁶⁶ However, some Japanese scholars worry that this revision may encourage large-scale platform enterprises that possess large amounts of data to lock up the data.⁶⁷ To alleviate the shrinkage of data flow and use caused by overly restrictive norms and rules, Japan insists on imposing restrictions to the minimum extent necessary in relation to specific uses of shared data of limited access.⁶⁸

Legislative approach: the Chinese solution for big data protection

There is some debate about how China should approach Big Data protection. For instance, some Chinese scholars claim that data producers' rights should be granted to those who

can produce or mine valuable data, taking inspiration from the EU's legislative framework.⁶⁹

While others believe that another viable option to protect Big Data collections is to incorporate exclusive rights into the framework of intellectual property law,⁷⁰ this may lead to excessive restrictions on the circulation and use of data. Long-term restrictions brought about by exclusive rights may stymie other uses of data that need not be restricted, especially when it does not concern the right-holder's competitive interest. Therefore, Big Data protection should not be achieved by granting persons exclusive rights.

Nevertheless, data theft by competitors should be subject to regulation.⁷¹ Scholars advocating the intellectual property rights approach observe that infringers almost always compete directly with obligees when the protection of Big Data is only limited to the right to prevent public dissemination. This is where anti-unfair competition law can operate because the public dissemination of data by others may amount to unfair competition.⁷²

A legislative discussion for Big Data protection requires consideration of the impact of having a monopoly of facts and information on competition policy, as well as academic research on ways to avoid excessive negative effects. Caution is advisable in the creation of new legislation. As we have learnt from the EU example, the protection afforded by the new EU *sui generis* right is unduly broad and has impeded fair competition. Indeed, the European Court of Justice had to come up with a restrictive interpretation to limit its scope of application. At the same time, the European Commission could unfortunately only voice its disapproval of the *sui generis* right in its report.⁷³

Thus, some scholars believe that legislation on the protection of data needs to consider the competitive situation of the market, and the scope of data protection should be delineated based on an examination of various elements such as the homogeneity of the market, the manner and purpose of use, and the relationship between the parties.⁷⁴ They suggest that any legislation on Big Data can be drafted with the aim of regulating competitive behaviour, which finds precedent in other legal systems such as Japan. This mode of regulation satisfies the need to fully explore and utilise Big Data in the era of the Internet of Things and Artificial Intelligence, avoids monopolistic behaviour of data enterprises, and balances data protection and circulation, making it an ideal mode for Big Data

⁵⁹ 不正競争防止法等の一部を改正する法律(平成30年5月30日法律第33号).(Law to Partially Revise the Anti-Unfair Competition Law, 30 May 2018, No. 33.)

⁶⁰ See Japanese Anti-Unfair Competition Law, Article 2, Paragraph 7. See also Liu Ying and Sui Ji Gang, "On the Adding of "Data Provision Restriction" Provision to Japanese Legislation Concerning Big Data and the Enlightenment to China", *Intellectual Property*, No. 4, 2019, pp. 93-94.

⁶¹ See Japanese Anti-Unfair Competition Law, Article 2, Paragraph 1, Item 11.

⁶² *Ibid*, Paragraph 1, Item 14.

⁶³ *Ibid*, Paragraph 1, Item 12, 13, 14, 15.

⁶⁴ *Ibid*, Article 3, Paragraph 1, 2.

⁶⁵ *Ibid*, Article 4.

⁶⁶ Some Chinese scholars championed the introduction of data clause in China. See Huaiyin Zhang, Yanhong Lou, Kui Cai, Research on the Dilemma and Improvement of Legal Regulation for Unfair Competition Related to Corporate Data in China, *Computer Law & Security Review*, Vol. 42, 2021.

⁶⁷ See 山内貴博「平成30年改正不正競争防止法への実務的対応」*ジュリスト*1525号(2018年)26頁.(See Takahiro Yamauchi, "Practical Response to the 2018 Revised Unfair Competition Prevention Law", *Jurist*, No. 1525, 2018, p. 26.)

⁶⁸ See 泉恒希「ビッグデータの法的保護に関する一考察」*知的財産法政策学研究*58号(2021年)191頁.(See Kohki Izumi, "A Study of the Legal Protection of Big Data", *Intellectual Property Law and Policy Journal*, Vol. 58, 2021, p. 191.)

⁶⁹ See *Supra* note 31, p. 125.

⁷⁰ See *Supra* note 10, p. 23.

⁷¹ Josef Drexler, "Designing Competitive Markets for Industrial Data: Between Propertisation and Access", *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, Vol. 8, 2017, p. 269.

⁷² See *Supra* note 10, p. 22.

⁷³ See 蘆立順美「欧州データベース保護指令に関する動向」*日本工業所有権法学会年報*30号(2007年)245頁.(See Masami Ashidate, "The review of the Directive on database protection", *Annual of Industrial Property Law*, No. 30, 2007, p. 245.)

⁷⁴ National Research Council, "A Question of Balance", 1999, p. 71; F. W. Grosheide, "Symposium on Intellectual Property, Digital Technology & Electronic Commerce: Digital Copyright and Database Protection: Database Protection- The European Way", *Washington University Journal of Law & Policy*, Vol.8, 2002, p. 55.

protection.⁷⁵ Therefore, the Chinese solution for Big Data protection should be a regime that prohibits unfair competition amongst data enterprises, instead of an intellectual property regime that is focused on exclusive “rights”.

Conclusion

In an era where the digital economy is rising in prominence, Big Data embodies significant economic and strategic value, making it not only a critical enterprise asset but also a strategic resource for China. The exploitation of Big Data often requires enormous capital investment, but the current law has not yet provided sufficient protection for Big Data in the case of unauthorized misappropriation by third parties. With the accumulation of practical experience, the boundaries and paths of Big Data protection will become increasingly apparent, not excluding the possibility of protection under civil rights in some cases. However, the selection of a legislative solution for Big Data protection should proceed prudently with respect for the existing legal framework and the values underlying it. Meanwhile, the institutional design of Big Data must take a careful and sober analysis of two issues: first, whether a new legal regime is really necessary; second, whether such a regime will incur disproportionate negative impacts. This article attempts to preliminarily explore the path of Big Data protection from a behavioural regulation perspective, with the hope of contributing to the upcoming task of building market rules for data in China and beyond.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Funding sources

Shanghai Pujiang Talent Program, “Innovation and Competition: Identification of Unfair Competition Behaviour in the Internet Age” (18PJC005), the Program for Professors of Special Appointment (Eastern Scholar) at [Shanghai Institutions of Higher Learning](#).

Acknowledgment

The authors wish to thank the following: Chong Fu Shan, Ge Jiangqiu, Li Ruoqing, Seah Ding Hang, Shi Hongpeng, Yuan Guohe.

Data availability

No data was used for the research described in the article.

⁷⁵ See Li Yang, “Review on Japan’s Unfair Competition Law Model of Data Protection”, *Journal of Political Science and Law*, No. 4, 2001, p. 76.