

Legal Issues Concerning Intellectual Property Rights in Enterprise Data in the Digital Economy

¹Qin Ran, ²Liu Yutong and ³Luan Minzhu,

^{1,2,3}Students, School of Law, Beijing University of Civil Engineering and Architecture, Beijing, China

Abstract—Against the backdrop of the digital economy, enterprise data has become a core factor of production across industries. At present, China's intellectual-property-related laws, including the Copyright Law and the Anti-Unfair Competition Law, still lack targeted regulatory mechanisms for protecting enterprise data. Using literature review, comparative analysis, and case analysis, this study systematically examines the nature, types, and legal attributes of enterprise data and identifies the limitations of applying traditional copyright, trade-secret, and patent regimes to such data. Drawing on representative disputes such as the Taobao Business Advisor case and the unfair-competition dispute involving the Shuabao app, the study further demonstrates four practical difficulties: unclear ownership, difficult transactions, frequent infringement, and impeded rights enforcement. Through a comparative analysis of the legislation and judicial implementation of the European Union's data-protection and database sui generis rights regime and the U.S. database protection regime, this article builds on the 'separation of three rights' policy in the Twenty Measures for Data. It proposes tiered rules for establishing rights in raw, processed, and derivative data; constructs a composite protection model comprising residual protection under the Anti-Unfair Competition Law, empowerment through a new sui generis data right, and blockchain-based registration and evidence preservation; and develops supporting mechanisms involving national legislation, data archiving, enterprise risk control, and the balancing of stakeholder interests. The findings provide a theoretical reference for the market-oriented allocation of data factors, the lawful and compliant use of enterprise data, and fair and consistent adjudication, thereby contributing to the improvement of China's digital rule-of-law framework.

Keywords—Digital economy; Enterprise data; Data intellectual property rights; Tiered rights determination; Data factor market

I. INTRODUCTION

A. Research Background

The digital economy has been elevated to a national strategy, and market-oriented reform of data factors continues to advance. As the digital economy expands and its share and importance in the market continue to grow, however, related legal loopholes and risks have emerged. These include the theft of enterprise data through web crawlers, illegal data resale, and a year-on-year increase in disputes over the ownership of platform data. Current law provides no clear means of resolving or penalizing these infringements. Articles 126 and 127 of the Civil Code contain only general provisions and establish no specific rules on intellectual property rights in data. Traditional intellectual property regimes cannot cover massive collections of enterprise data, resulting in the practical contradictions of difficult rights determination, difficult enforcement, and difficult transactions.

B. Research Significance

1. Theoretical Significance: This study helps fill gaps in the theory of intellectual property rights in data, improve the intellectual property law framework, and respond to scholarly disputes concerning data ownership and usufructuary rights in data.
2. Practical Significance: The study offers proposals for enterprise data compliance, judicial adjudication of data-infringement disputes, and national data legislation. It may also reduce enterprise data transaction costs and foster a sound digital business environment.

II. THEORETICAL FOUNDATIONS OF INTELLECTUAL PROPERTY RIGHTS IN ENTERPRISE DATA

Enterprise data refers to collections of information formed through lawful collection, cleaning, and processing. It includes customer information, industry operational data, and algorithmically derived data products and can be divided into three categories. Raw data includes basic user information and original industry data. Processed data consists of data clusters that enterprises have screened, labeled, and integrated. Derivative data consists of commercially valuable data products created through intelligent algorithmic processing and synthesis. Enterprise data is intangible, fluid, and collective and carries attributes of both personality rights and property rights, distinguishing it from traditional works, technologies, and patents.

Rights and interests in enterprise data are reflected to some extent in foundational laws such as the Civil Code, Copyright Law, Anti-Unfair Competition Law, and Data Security Law. Nevertheless, theoretical controversy remains over whether enterprise data can constitute an independent object of intellectual property rights, whether traditional intellectual property protection pathways can be applied to it, and whether original compilations of data qualify for copyright protection.

III. CURRENT STATUS AND CHALLENGES OF LEGAL PROTECTION FOR INTELLECTUAL PROPERTY RIGHTS IN ENTERPRISE DATA IN CHINA

A. Deficiencies in the Adaptation of Existing Intellectual Property Regimes

Traditional intellectual property regimes use originality, novelty, and distinctiveness as their core requirements for rights determination and are designed for static, finalized intellectual achievements with clearly identifiable creators. By contrast, the core value of enterprise data lies in its comprehensiveness, objectivity, continuous dynamic updating, formation through multi-party participation, and absence of a single creator.

Traditional copyright protects only original intellectual creations. Enterprises, however, invest substantial capital, compliance costs, and operational labor in collecting, acquiring, and storing raw data to form collective datasets.

Although such datasets may lack originality, they possess objective commercial and competitive value. Permitting unrestricted data scraping would weaken incentives for enterprise innovation and violate the principle of fairness. If underlying data receives no legal protection, enterprises will be reluctant to make long-term investments in collecting and processing data, thereby impeding the market-oriented circulation of data factors and innovation in digital industries. To encourage sustained data-factor development and safeguard the sustainable growth of the digital economy, a new mechanism for determining intellectual property rights in enterprise data must account for substantive investment, lawful possession, and the maintenance of market competition.

Trade-secret protection requires secrecy and therefore constitutes a closed, static form of protection. Modern enterprise data assets, including raw, processed, and derivative data, often need to be traded, pledged, and valued, which may require disclosure and broad circulation. Once disclosed externally, the data loses its secrecy. Trade-secret protection is consequently ill-suited to standardized and repeated data transactions and cannot support the market-oriented operation of data assets as a whole. Moreover, enterprise data lacks a property-like system of public notice, rights registration, and confirmation. Standard intellectual property transactions such as title transfer, pledge financing, and asset securitization therefore cannot be completed effectively, resulting in high transaction costs, ambiguous rights boundaries, and valuation difficulties. Bulk web scraping and algorithm theft are highly concealed in the digital era, yet a trade-secret claimant must prove secrecy, value, and protective measures. Evidence collection is therefore difficult, and comprehensive disclosure can leave the claimant without an effective remedy.

Patent law protects technical solutions rather than enterprise data itself. Raw data is objective information and cannot be patented, while the novelty and inventiveness thresholds for patents are high. The patent model of exchanging disclosure for protection is also inconsistent with the need to preserve confidentiality during data circulation. Patent remedies constrain technical copying but have difficulty regulating data scraping and resale, cannot determine ownership of data, and are poorly adapted to data transactions. Patent law can therefore play only a supporting role.

B. Practical Challenges Revealed by the Taobao Business Advisor Data Dispute

In the Taobao Business Advisor data dispute, the court ruled solely under the Anti-Unfair Competition Law and did not recognize independent intellectual property rights in enterprise data. The case demonstrates that China's current legal framework lacks specialized rules for protecting property interests in data and clear, consistent standards for determining infringement. Four principal difficulties arise.

1. **Ambiguous Rights Boundaries:** Existing law does not clearly specify whether rights can be established in raw enterprise data, what standards apply, or how rights models should differ.
2. **Frequent Data Infringement:** Common conduct includes using web crawlers to collect data for profit, disclosure and resale of data by departing employees, and copying data to create derivative products in the same industry.
3. **Insufficient Judicial Protection:** Courts face difficulty determining whether an enterprise has made a substantive labor or capital contribution, distinguishing reasonable data use from malicious scraping, and quantifying property losses under uniform rules.

4. **Constraints on Data-Factor Circulation:** The absence of standards for compliant enterprise data circulation increases transaction costs and risks and hinders the development of the data factor market.

IV. INTERNATIONAL EXPERIENCE AND DOMESTIC APPLICATION OF ENTERPRISE DATA PROTECTION

Countries around the world have developed diverse mechanisms for protecting intellectual property interests in enterprise data, offering valuable references for China's institutional design.

1. **European Union:** The EU combines the General Data Protection Regulation (GDPR) with a database sui generis rights regime. It strengthens data protection while seeking to realize the circulation value of data and grants limited exclusive rights in databases formed through enterprise selection and processing. Nevertheless, the model significantly restricts the commercial circulation and factor-use efficiency of enterprise data. Its emphasis on rights protection may consequently inhibit development of the data industry.
2. **United States:** The United States employs a dual mechanism involving copyright and special database protection, with property-interest protection as its core rationale. It focuses on safeguarding capital and labor invested in data collection, organization, and operation and protecting investor returns, thereby adapting well to market-based data competition. However, its strong emphasis on enterprise property interests insufficiently balances personal information rights and may facilitate data misuse and privacy infringement.
3. **Japan and South Korea:** Both countries have actively piloted data-rights registration. Their standardized registration systems clarify rights in data and provide practical experience for compliant data circulation.

Given the current development of China's data factor market, neither the European nor the American model can be copied wholesale. China's intellectual property system for data must be grounded in local needs, adhere to the principle of balancing interests, account for the rapid growth of the domestic digital economy, and draw selectively on international experience to build a data-rights protection system with Chinese characteristics. Its implementation must balance the interests of data holders, users, and the public. A classified and tiered data-management system should apply differentiated protection strategies: core data involving national security, public interests, or personal privacy should receive strict protection, whereas reasonable development and use of ordinary commercial data should be encouraged. Data transaction rules and a secure, trustworthy circulation environment should also be improved. Legal and technical safeguards should jointly ensure security and controllability throughout circulation. Clear standards and legal liabilities for infringement, together with effective remedies for data rights holders, can foster a constructive interaction between data protection and use.

V. INSTITUTIONAL DESIGN FOR PROTECTING INTELLECTUAL PROPERTY RIGHTS IN ENTERPRISE DATA

A. Tiered Rules for Determining Rights in Enterprise Data

1. **Raw Data:** Raw data is naturally linked to the personal information rights of natural persons and has strong personality attributes. Rights determination should therefore prioritize users' personality rights. An enterprise may obtain only a limited, non-exclusive, and temporary

license to use raw data through user authorization or lawful public channels and may not enjoy full rights of possession, disposition, or independent transfer. Where several parties collect data jointly, each party may share only read access, and no party may package and trade the raw data independently. Reasonable-use exceptions should permit limited use of publicly available raw data for public-interest purposes such as market analysis and industry research.

2. **Processed Data:** Processed data is generated through an enterprise's investment of labor and capital in basic processing but has not undergone deep algorithmic reconstruction or become an independently marketable product. The enterprise should enjoy a complete right to process and use such data, including internal deployment, external licensing, and limited transfer. Exclusive use should be tied to continuous operation and maintenance; prolonged discontinuation of updating and maintenance should weaken exclusivity to preserve the circulation of foundational industry data.
3. **Derivative Data Products:** Derivative data products are independently valuable and directly tradable products formed when an enterprise uses algorithmic models and computing power to conduct deep modeling, computation, and visual reconstruction of processed data. No person may copy, move, alter, or resell a complete derivative data product without authorization. For derivative data generated by artificial intelligence, the algorithm developer should be the holder of the sui generis right. The rules must clearly distinguish simple use of a tool from substantive investment in deep algorithmic processing to prevent the use of algorithms to misappropriate others' data products.
4. **Whole-Process Control:** The three-tier rights framework should be supported by an online tiered registration system. Enterprises may separately file processed datasets and derivative data products with a unified national data intellectual property registration platform. Raw data should be recorded only for the scope of authorized use and not registered as owned property. Separate records for all three tiers would establish a complete rights-management chain covering collection, processing, output, circulation, and enforcement and provide legally recognized evidence for transactions and litigation.

B. A Composite Protection Model: Residual Protection, Sui Generis Data Rights, and Blockchain Evidence Preservation

1) Residual Protection Under the Anti-Unfair Competition Law

This mechanism applies to ordinary datasets that lack the originality required for intellectual property protection and do not qualify for a sui generis data right, including de-identified raw databases and basic processed databases. The general and internet-specific provisions of the Anti-Unfair Competition Law should regulate boundary-crossing web scraping, employee data resale, copying of basic data by competitors, and large-scale use beyond the scope of authorization. The mechanism should balance data circulation by small and medium-sized enterprises against protection for platform investment and prevent monopolization of the data market.

2) Empowerment Through a New Sui Generis Data Right

The new sui generis data right would be an independent form of intellectual property distinct from copyright, patent rights, and trade secrets. Upon completing online registration, an enterprise would obtain statutory exclusivity, including rights of exclusive use, paid licensing, profitable transfer, and independent intellectual property litigation. Infringement would support punitive damages, raising the cost of

infringement and addressing the inadequacy of current awards. For AI-generated derivative data, the algorithm developer would be the holder of the right. A distinction between tool use and substantive investment in deep processing would help prevent gray-area misappropriation of data products.

3) Blockchain-Based Registration and Evidence Preservation

Blockchain-based registration and evidence preservation should extend across proof of rights, infringement evidence collection, and data transactions to address evidentiary difficulties in litigation. In a rights dispute, on-chain records could be admitted directly as legally recognized evidence without requiring an enterprise to submit additional development logs or investment records. For web scraping and data theft, platform access and retrieval records could be placed on-chain in real time, conclusively identifying the infringer, time, and scale of infringement and enabling accurate loss quantification. Deep integration with notarization and judicial systems would allow courts and notary offices to retrieve on-chain evidence directly and substantially simplify enforcement.

C. Improvement of Judicial Adjudication and Remedies

Implementation of the composite data protection model requires detailed and uniform adjudication rules and tiered remedies. These are necessary to address inconsistent judicial standards, ambiguous compensation, and high evidentiary thresholds and to standardize the judicial protection of data interests. A uniform three-step infringement test should apply to all data-infringement cases.

1. **Legality Review:** Courts should verify the lawfulness of data collection, user authorization, and de-identification. Claims based on unlawfully sourced data should be dismissed to prevent enforcement of illegal data interests.
2. **Substantive Investment Assessment:** Courts should consider an enterprise's financial, labor, algorithm-development, and maintenance investments, distinguish independently processed data from public raw data, and deny protection to foundational data created without substantive investment.
3. **Competitive Harm Assessment:** Courts should consider commercial value, scale of infringement, and subjective malice to distinguish reasonable retrieval from predatory scraping and unify adjudication standards nationwide.

Tiered compensation standards should be refined through representative cases. Raw and basic processed data protected under the Anti-Unfair Competition Law should receive compensation based on the rights holder's actual loss or the infringer's unlawful gains, with statutory damages available when neither can be quantified. Punitive damages should apply to derivative data products protected by a sui generis data right. Malicious or repeated infringement could support an award of one to five times the actual loss, and graded awards should distinguish fragmented appropriation from resale of a complete product.

Diversified evidence rules should reduce enforcement thresholds. Electronic notarization should be conducted in advance, platform-log retention should be standardized, and judicial authorities should be able to retrieve back-end records directly. For large-scale crawler infringement, the burden of proof should shift to the accused party to establish legality; absent valid evidence, infringement should be found. These measures would simplify procedures and improve enforcement efficiency.

VI. SUPPORTING MECHANISMS FOR PROTECTING INTELLECTUAL PROPERTY RIGHTS IN ENTERPRISE DATA

The implementation of tiered rights determination and composite protection depends on four coordinated systems: national institutions, registration and evidence preservation, multi-party coordination, and enterprise risk control. Together, they can address fragmented governance rules, inconsistent standards, enforcement difficulties, and imbalances of interests.

A. National Legislation and Specialized Supervision

China's current rules on intellectual property rights in data are fragmented, lack specialized legislation, and rely excessively on policies with limited binding force. Special legislation should therefore be advanced on the basis of the 'separation of three rights' approach, clarifying the boundaries of the three data tiers and the legal status of the new sui generis data right. Reasonable-use and inclusive-access provisions should be added to prevent platform data monopolies. The National Data Administration should coordinate reform pilots, unify national rights-determination and transaction standards, and adopt support policies that reduce compliance costs for small and medium-sized enterprises. A specialized regulator should coordinate multi-agency enforcement, conduct market inspections and compliance warnings, and establish cross-regional enforcement mechanisms to remove regulatory barriers.

B. A Unified National Registration and Blockchain Evidence System

The state should coordinate an online registration platform with differentiated filing and rights-determination modules for raw, processed, and derivative data and issue universally recognized electronic certificates. A consortium blockchain should make registration information, transaction contracts, and authorization records immutable and support filings for iterative data updates. Connections with courts, notary offices, and data exchanges would enable cross-institutional recognition and one-click retrieval of evidence, simplify enforcement, eliminate uncertainty over rights in transactions, and stimulate circulation in the data factor market.

C. Coordination of Individual, Enterprise, and Government Interests

The personality rights of natural persons should receive priority, and data should be de-identified before rights can be registered, thereby safeguarding personal information. Reasonable access to foundational data should be moderately expanded to support non-commercial retrieval by researchers and micro, small, and medium-sized enterprises. Routine antitrust supervision should prevent data monopolies. Public data should be owned by the state, with enterprises enjoying only processing and use rights, while a public-interest exception should permit uses required for emergency governance and macroeconomic regulation.

D. Internal Enterprise Compliance and Risk Control

Enterprises should implement three-tier data controls, refine access permissions and operational records, strictly control the circulation of raw data, and retain research and development logs for derivative data. Internal and external authorization and approval procedures should be standardized, and departing employees' access should be revoked under enforceable confidentiality mechanisms. Web-crawler monitoring and bulk-download warning systems should automatically intercept infringement and preserve on-chain evidence. Routine compliance self-inspections and third-party

audits would prevent risks and ownership disputes at their source and provide whole-process protection for data interests.

CONCLUSION AND OUTLOOK

China's traditional intellectual property regimes cannot adequately meet the protection needs of enterprise data and create multiple difficulties in rights determination, enforcement, and transactions. Drawing on international experience, tiered rights determination and a diversified composite protection model offer the institutional path best suited to the digital economy. As the digital economy continues to develop, data infringement is likely to increase. China must improve its laws and specify effective sanctions to regulate the data environment and strengthen the legal system. Legislation, adjudication, registration, and enterprise risk-control mechanisms are all indispensable. As market-oriented reform of data factors advances, specialized regulations on intellectual property rights in data are likely to emerge, blockchain registration will be implemented more broadly, and stable and uniform judicial standards for data infringement will take shape, ultimately producing a more complete legal framework.

References

- [1] H. Xu, "Web scraping of platform data for profit results in compensation exceeding RMB 230,000," *Beijing Daily*, Jun. 9, 2026.
- [2] M. Zhou, G. Liu, and T. Xie, "Compliance pathways for the determination and circulation of enterprise data assets from the perspective of civil and commercial law," *China Business*, no. 11, 2026.
- [3] L. Ji, "Differentiated calculation rules for damages arising from enterprise data infringement," *Governance Studies*, no. 3, 2026.
- [4] X. Han, "A normative mechanism for determining rights in enterprise data," *Seeker*, no. 2, 2026.