Simen Owen Academic Proceedings Series

3rd International Conference on Education, Environment, Arts and Social Science (EEAS 2026)

Article

Application Dilemmas and Institutional Improvements of the Crime of Infringing Citizens' Personal Information in the Context of Big Data

Jiaqi Qu [1,*]

[1] Criminal Justice School, Zhongnan University of Economics and Law, Wuhan, China

* Correspondence: Jiaqi Qu, Criminal Justice School, Zhongnan University of Economics and Law, Wuhan, China

Abstract: Under the conditions of big data, the crime of infringing citizens' personal information has become significantly more difficult to apply in judicial practice because personal information is increasingly collected, aggregated, transferred, and reused across diverse platforms, institutions, and complex data-processing chains. This paper systematically examines the application dilemmas of this specific crime within the framework of Chinese criminal law, focusing primarily on the inherent tension between strict legal certainty and the dynamic, complex realities of modern data circulation. It argues that while existing quantity-based standards remain necessary for baseline adjudication, they are fundamentally insufficient to address nuanced cases involving highly sensitive information, authorised access followed by unauthorised secondary use, malicious insider leakage, sophisticated technical circumvention, doxxing databases, and severe downstream fraud. By adopting a rigorous methodology encompassing normative analysis, structured case observation, and limited case coding, this paper critically analyses typical cases recently released by the Supreme People's Court. These include medical data leakage, railway passenger-information sale, illegal acquisition of education records, doxxing through social-engineering databases, and fraud based on HPV vaccine-reservation information. The study ultimately finds that judicial reasoning must urgently move beyond a single quantity standard and adopt a comprehensive, multi-dimensional assessment based on information type, data source, conduct method, actor responsibility, and harmful consequence. It further proposes actionable institutional improvements concerning personal-information identification, seriousness assessment, administrative-criminal boundary clarification, evidence review, and corporate compliance governance.

Keywords: big data; criminal law; data compliance; judicial application; personal information; information infringement

Received: 12 April 2026
Revised: 18 May 2026
Accepted: 28 May 2026
Published: 04 June 2026

1. Introduction

The development of big data has changed the way citizens' personal information is collected, stored, transferred, and used. In traditional cases, infringed information often appeared as simple identifiers, such as names, telephone numbers, addresses, or identity-card numbers [1]. Under big-data conditions, however, personal information is increasingly embedded in platform databases, transaction records, location traces, medical records, educational records, and behavioural profiles. Data collected for legitimate purposes may later become objects of illegal sale, provision, excessive use, or secondary exploitation. Therefore, the crime of infringing citizens' personal information now concerns not only individual privacy, but also data security, platform governance, and the order of information circulation.

In Chinese criminal law, Article 253A of the Criminal Law and relevant judicial interpretations provide the basic normative framework for punishing the illegal acquisition, sale, or provision of citizens' personal information. These rules have helped address large-scale information leakage and data black-market transactions [2]. Nevertheless, judicial practice still faces difficulties in the big-data context. The scope of "citizens' personal information" has become more complex, especially when data are aggregated, de-identified, inferred, or combined with other datasets. The quantity standard, although necessary for legal certainty, cannot fully reflect differences in sensitivity, source, usage scenario, and social harm. The boundary between administrative illegality and criminal liability also remains unclear in cases involving platforms, outsourcing service providers, and internal employees.

Existing research has contributed to the interpretation of the offence, especially regarding protected legal interests, conduct types, and sentencing standards. However, some studies remain focused on abstract doctrinal analysis, while others describe judicial decisions without offering a structured method for comparing different cases [2]. In particular, further analysis is needed for mixed-data cases, authorised access followed by unauthorised use, and the relationship between upstream information infringement and downstream crimes such as telecom fraud or online harassment.

This paper examines the application dilemmas of the crime of infringing citizens' personal information under big-data conditions and proposes practical improvement paths. Its contributions are threefold. First, it proposes a five-dimensional analytical framework based on information quantity, information type, data source, conduct method, and harmful consequence. Second, it develops a classification approach for mixed personal-information cases, so that quantity standards can be considered together with sensitivity and scenario-based risks [1]. Third, it explores the coordination between criminal regulation, administrative enforcement, and enterprise compliance, with the aim of clarifying the boundary between general data-compliance violations and criminal offences.

Methodologically, this paper adopts normative analysis, case analysis, and limited empirical observation of judicial decisions. It first clarifies the legal structure of the offence, then identifies major difficulties in judicial application, and finally proposes improvement measures concerning information identification, seriousness assessment, evidence review, and compliance governance. The study seeks to provide a more stable and explainable approach to criminal-law application while respecting the modest and supplementary role of criminal law [3].

2. Legal Norms and Practical Development of the Crime

2.1. Normative Basis of the Crime

The crime of infringing citizens' personal information is mainly regulated by Article 253A of the Criminal Law and related judicial interpretations. Its establishment reflects the criminal law's response to the increasing social risks caused by illegal collection, sale, provision, and use of personal information. In the big-data context, personal information is no longer merely an object of private interest [4]. Once it is illegally circulated, it may affect personal security, property security, platform order, and the broader governance of data flows. Therefore, the protected legal interest of this crime should be understood as a compound interest: it includes individual rights to personal information, and also the public order of lawful information circulation.

In addition to criminal law, the Personal Information Protection Law, the Data Security Law, and relevant administrative regulations provide important background norms. These rules clarify principles such as legality, legitimacy, necessity, purpose limitation, and security protection. Although violation of these rules does not automatically lead to criminal liability, they help determine whether the acquisition, provision, or use of personal information is unlawful [4]. In this sense, the crime must be understood within a broader system of data governance.

2.2. Legal Elements of the Crime

The object of this crime is citizens' personal information. In judicial application, personal information generally refers to information that can identify a specific natural person, either independently or in combination with other information. Traditional identifiers, such as names, identity-card numbers, telephone numbers, and addresses, are relatively easy to identify. However, under big-data conditions, location records, transaction data, medical information, educational records, account data, device identifiers, and behavioral profiles may also have identifying value [5].

The conduct elements mainly include illegal acquisition, sale, and provision of citizens' personal information. "Illegal acquisition" may include purchasing, stealing, fraudulently obtaining, technically intruding into systems, or exceeding authorized access. "Sale" and "provision" focus on the transfer of information to others, whether for direct profit or other improper purposes. The subjective element is generally intentional [5]. The actor must know, or at least clearly recognize, that the information is citizens' personal information and that the acquisition, sale, or provision violates relevant rules.

The crime also requires the conduct to reach a certain level of seriousness [6]. Quantity is an important factor, but it should not be the only factor. Information type, sensitivity, source, actor identity, illegal gains, and harmful consequences should also be considered when determining whether the circumstances are serious or especially serious.

2.3. New Characteristics under Big Data

Big data has transformed the practical dynamics of this crime in several ways [7]. The scale of infringement has significantly expanded, with a single leakage potentially involving thousands or even millions of records. Additionally, the types of information have become increasingly complex, as various categories of data can be combined to create a more precise profile of an individual. Furthermore, the sources of infringement have become more concealed, often involving insiders, outsourced service providers, platform employees, or technical personnel with access to data systems.

Personal information crimes are frequently linked to downstream offenses such as telecom fraud, online harassment, extortion, and illegal marketing. In these instances, the societal harm caused by information infringement extends beyond the initial data transfer [8]. The information may be repeatedly resold, recombined, and reused, resulting in diffuse and challenging-to-trace consequences.

2.4. Tension between Data Use and Criminal Regulation

The application of this crime must balance data use and criminal regulation [9]. On the one hand, the digital economy depends on lawful data circulation, and not every compliance defect should be treated as a criminal offence. Over-expansion of criminal liability may create uncertainty for normal business activities. On the other hand, large-scale illegal circulation of personal information can seriously undermine personal security and social trust. Criminal law should therefore intervene where the conduct shows clear illegality, substantial scale, profit-seeking purpose, serious risk, or actual harm.

Accordingly, the key issue is not whether personal information should be protected, but how to draw a reasonable boundary between administrative violation, civil liability, and criminal punishment. This boundary should be based on the nature of the information, the way it is obtained or provided, the actor's role, and the actual or realistic risk caused by the conduct [10].

3. Application Dilemmas in Judicial Practice

Under big-data conditions, the application of the crime of infringing citizens' personal information faces difficulties not only due to the increasing scale of data but also because of the evolving forms of data circulation [1]. Personal information may initially be lawfully collected but later used, transferred, or combined beyond its authorized purpose. Judicial practice must address five interconnected dilemmas: definition, quantity and sensitivity, administrative-criminal boundary, actor responsibility, and harm

causation. As illustrated in Figure 1, these dilemmas primarily involve defining personal information, assessing its quantity and sensitivity, delineating the boundary between administrative and criminal liability, identifying actor responsibility, and proving harm and causation.

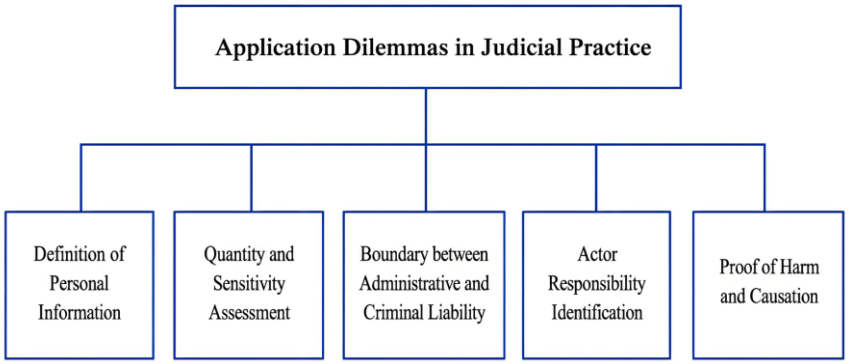

Figure 1. Structure of Application Dilemmas in Judicial Practice

3.1. Difficulty in Defining Citizens' Personal Information

The first difficulty lies in defining "citizens' personal information." Traditional identifiers, such as names, telephone numbers, addresses, and identity-card numbers, are relatively easy to recognize. However, big-data cases often involve indirect or contextual information, such as account data, location traces, transaction records, medical records, education records, and behavioral profiles. Such data may not identify a person independently but may become identifiable when combined with other information [11].

The education-record case released by the Supreme People's Court illustrates this issue. The defendants bypassed verification on an official education-information platform and downloaded and sold electronic academic-registration records [12]. This case shows that education data may affect employment, qualification review, reputation, and identity verification. Therefore, the identification of personal information should depend not only on the data field itself but also on identifiability, use scenario, and realistic risk.

3.2. Difficulty in Determining Quantity and Sensitivity

The second difficulty concerns the relationship between quantity and sensitivity. Quantity standards are necessary because they provide relatively clear criteria [2]. However, quantity alone cannot fully reflect the seriousness of big-data infringement. A small amount of medical, financial, travel, or education information may create greater risk than a larger amount of ordinary contact information.

The medical outsourcing case is representative. A software company responsible for maintaining an online hospital-registration system illegally obtained patient-registration information during service provision. The case shows that courts must consider not only the number of records, but also deduplication, medical sensitivity, and whether the actor abused a service relationship [13].

The "doxxing" database case further reveals the limits of a purely quantity-based approach. The defendants used a large-scale social-engineering database that could support targeted exposure, harassment, insult, and online violence. In such cases, harm arises from both data scale and accessibility, and extends to repeated search, disclosure, and secondary misuse.

3.3. Difficulty in Distinguishing Administrative Illegality from Criminal Liability

The third difficulty lies in defining the boundary between administrative violations and criminal offenses. Enterprises may breach compliance duties through excessive data collection, insufficient consent, weak security management, or improper outsourcing. However, not all compliance defects should be classified as crimes [7]. Criminal liability

should be reserved for actions characterized by clear illegality, substantial scale, profit-driven motives, significant risk, or actual harm.

This issue is particularly evident in platform and outsourcing scenarios. In a medical outsourcing case, the actor initially gained system access due to a service relationship but later misused that access beyond the authorized purpose. This demonstrates that lawful access at the outset does not necessarily justify subsequent acquisition, storage, or transfer. The critical factor is whether the actor remained within the authorized scope and lawful purpose.

3.4. Difficulty in Identifying Actor Responsibility

The fourth difficulty concerns responsibility allocation. Big-data crime may involve internal employees, outsourcing companies, platform personnel, data brokers, purchasers, and downstream users. Their responsibility should be differentiated according to role, knowledge, profit, and contribution [2].

The railway passenger-information case shows the special risk of internal personnel. The defendant used occupational access to query passengers' travel information and sold it for profit. Such conduct is more harmful than ordinary external acquisition because it abuses institutional trust and weakens data-management security. However, organizers, technical assistants, sellers, and purchasers should not automatically bear the same degree of liability.

3.5. Difficulty in Proving Harm and Causation

The fifth difficulty is proving harm and causation. Once personal information enters the black market, it may be resold, recombined, and used by different actors [8]. It is often challenging to establish that a specific loss was directly caused by one particular leakage.

The HPV vaccine-reservation case provides a clearer causal chain. The defendants obtained reservation information through technical means and then used phishing websites and text messages to commit fraud. In such cases, the connection between information infringement and downstream crime is relatively direct. By contrast, in ordinary data-resale cases, courts may need to assess concrete and foreseeable risks rather than rely solely on actual losses. Harm assessment should therefore combine actual consequences with realistic risks, while avoiding an overbroad use of abstract danger.

In summary, these dilemmas arise from the interaction between legal standards and big-data practices. Judicial reasoning should move beyond a single quantity standard and adopt a more comprehensive assessment of information type, data source, actor role, conduct method, and harmful consequence [5, 8].

4. Structured Observation of Typical Cases

This chapter conducts a structured observation of typical cases released by the Supreme People's Court, including medical data leakage by an outsourcing company, railway travel-information leakage by an internal employee, illegal acquisition of education records, "doxxing" through a social-engineering database, and fraud based on HPV vaccine-reservation information [7]. These cases reflect several recurring patterns in current personal-information crimes: source leakage, insider abuse, technical circumvention, online secondary harm, and connection with downstream offences.

4.1. Corpus and Coding Method

This paper adopts a limited case-coding method. The purpose is not large-scale quantitative modeling, but to observe how courts identify key factors in different cases. Each case is coded according to information type, data source, conduct method, actor identity, illegal gains, downstream use, and judicial result. Table 1 outlines the coding items used in this chapter.

Table 1. Coding Items for Judicial Case Analysis

Coding Item	Meaning	Function in Analysis
Information type	Medical, travel, education, account, reservation, or identity information	Assess sensitivity
Information quantity	Number of records after verification or deduplication	Assess seriousness
Data source	Hospital system, railway system, official platform, website, or underground database	Identify leakage channel
Conduct method	Illegal acquisition, sale, provision, intrusion, or unauthorised use	Identify conduct
Actor identity	Outsourcing company, insider, technical actor, data broker, or downstream user	Assess responsibility
Downstream use	Fraud, doxxing, illegal sale, harassment, or other misuse	Assess social harm
Judicial result	Conviction, sentence, fine, public-interest remedy, or combined punishment	Observe reasoning pattern

4.2. Distribution of Case Types

The five typical cases show that the crime is no longer limited to the simple sale of contact information. The first scenario is source leakage, represented by the medical outsourcing case. The defendant company maintained a hospital-registration system but secretly collected patient-registration information and imported it into a self-built database. After deduplication, the data amounted to 2,878,070 records. The case demonstrates that service access does not justify collection beyond the authorized purpose.

The second scenario is insider abuse, represented by the railway passenger-information case. The defendant used occupational access to query travel information, including train number, stations, seat information, and identity-document data, then sold it for profit. The illegal gain reached approximately 190,000 yuan. This highlights that internal personnel may create higher risks due to their institutional access to sensitive systems.

The third scenario is technical circumvention, represented by the education-record case. The defendants used false identity materials, mobile numbers, and verification bypass methods to obtain academic-registration records from an official platform. This underscores that education data may influence employment, qualification reviews, and social credit, and should not be treated as ordinary low-risk data. Table 2 further summarizes these typical case types and their corresponding legal significance.

Table 2. Types of Typical Cases and Their Legal Significance

Case Type	Representative Case	Main Legal Issue
Source leakage	Medical outsourcing case	Abuse of service access and unauthorised data collection
Insider abuse	Railway passenger-information case	Occupational convenience and breach of institutional trust
Technical circumvention	Education-record case	Bypassing platform verification and illegally obtaining official records

Doxxing database	Social-engineering database case	Online exposure, harassment, and secondary misuse of aggregated data
Downstream fraud	HPV vaccine-reservation case	Link between information infringement and targeted fraud

4.3. Judicial Reasoning Patterns

From these cases, courts mainly rely on five factors: information quantity, information sensitivity, illegality of the method, actor role, and harmful consequence. Quantity remains important, but it is not the only basis. In the medical outsourcing case, the large number of records was relevant, but the medical nature of the data and the abuse of a service relationship were also decisive [4]. In the railway case, the court emphasized the defendant's occupational convenience and the sensitive nature of travel information. These factors show a shift from simple quantity assessment to contextual evaluation.

The "doxxing" case further confirms this trend. The defendants illegally obtained hundreds of millions of records and built a social-engineering database containing more than 170 million records. The database was accessed more than 100,000 times and was connected with targeted exposure, insult, harassment, and online violence. The seriousness of the case therefore derived from both scale and repeated secondary misuse.

4.4. Associated Crimes and Harm Assessment

The HPV vaccine-reservation case illustrates how the infringement of personal information can facilitate downstream fraud. The defendants embedded a Trojan program into a legitimate vaccine-reservation website, acquiring over 290,000 reservation records. They subsequently used phishing websites and text messages to defraud 51 victims of more than 580,000 yuan. The court imposed combined punishment for both fraud and the infringement of citizens' personal information. This approach is compelling, as the stolen information directly enabled targeted deception against identifiable victims.

These cases collectively highlight three key observations. First, courts place significant emphasis on sensitive information, particularly data related to medical, travel, education, and health-related reservations. Second, the identity of the actor is crucial: insiders and service providers face stricter scrutiny when they misuse authorized access. Third, the downstream use of information significantly influences legal evaluations, especially when it is employed for fraud, doxxing, harassment, or other related crimes. These findings support a comprehensive assessment framework based on the type of information, its source, the method of conduct, the actor's responsibility, and the resulting harm [12].

5. Improvement Path of Legal Application

The previous chapters demonstrate that the primary challenge lies in bridging the gap between general legal standards and the complexities of data practices [13]. Under big-data conditions, courts should not rely solely on the number of records but should comprehensively consider factors such as the type of information, the source of data, the method of conduct, the responsibility of the actor, and the resulting harmful consequences.

5.1. Improve the Standard for Identifying Personal Information

The identification of personal information should adhere to the standard of "identifiability plus contextual risk." Traditional identifiers, such as names, telephone numbers, addresses, and identity-card numbers, should be included. For device identifiers, account data, education records, transaction information, location traces, and behavioral profiles, courts should assess whether the data can identify a specific person independently or in combination with other information [3, 10].

However, criminal-law protection should not be expanded without limits [4]. If data have been effectively anonymized and cannot identify a specific person through reasonable technical means, they should not be easily classified as citizens' personal

information. The key is to differentiate genuine anonymization from superficial de-identification.

5.2. Establish a Comprehensive Seriousness Assessment Standard

The standard for "serious circumstances" should transition from a singular focus on quantity assessment to a more comprehensive risk evaluation. While quantity remains an important factor, it alone cannot fully capture the gravity of issues such as medical information leakage, the sale of railway travel information, doxxing databases, or vaccine-reservation information fraud.

Courts should evaluate seriousness based on five key factors: the quantity of information, the type of information, the source of the data, the method of conduct, and the resulting harmful consequences. Sensitive information, including medical, travel, financial, educational, and health-related reservation data, warrants closer scrutiny. Data obtained from internal systems, official platforms, or service relationships should also be subject to stricter evaluation, as such cases often involve an abuse of trust or authority.

In cases involving mixed data, courts should avoid mechanically aggregating all data. Instead, different categories should first be classified based on their sensitivity and thresholds, followed by a proportional evaluation.

5.3. Clarify the Boundary between Administrative and Criminal Liability

The distinction between administrative violations and criminal liability must be clearly defined. Personal information processing may breach administrative regulations through insufficient consent, excessive data collection, inadequate security measures, or improper outsourcing practices. Such breaches should primarily be addressed through administrative penalties, civil liability, rectification orders, and compliance oversight.

Criminal law should only apply when the actions exhibit a higher degree of illegality and social harm. Courts should assess whether the individual knowingly exceeded authorization, whether the actions involved large-scale data acquisition or transfer, whether illegal profits were sought, whether sensitive information was implicated, and whether actual harm or significant risk was present.

5.4. Strengthen Evidence Review Rules

Evidence review should be strengthened [5]. Data evidence may contain duplicates, invalid entries, incomplete records, or information that cannot identify a natural person. Courts should examine the legality of data extraction, the reliability of electronic evidence, the calculation method, and the deduplication process.

For cases involving downstream harm, such as fraud or online harassment, courts should analyze the causal chain between information infringement and subsequent misuse. Where direct causation is difficult to prove, concrete and foreseeable risks may be considered, but abstract assumptions should be avoided.

5.5. Improve Corporate Compliance and Preventive Governance

Criminal-law application should also be integrated with preventive governance. Many cases stem from inadequate internal controls, excessive employee access, insufficient log management, or lax outsourcing supervision. Enterprises handling significant amounts of personal information should implement classification systems, minimum-necessary access rules, regular permission reviews, access-log retention, data export approval processes, and internal accountability mechanisms.

For outsourcing scenarios, data controllers should clearly define processing purposes, data scope, storage periods, confidentiality obligations, and breach-response responsibilities within contracts. They should also conduct audits to ensure that service providers do not process data beyond their authorization.

In conclusion, enhancing the application of this crime requires a balanced approach. Criminal law should effectively safeguard citizens' personal information while remaining measured and precise. A comprehensive framework based on information type, quantity, source, conduct method, actor role, and harmful consequences can better address big-data risks while ensuring legal certainty and proportionality [10].

6. Conclusion

This paper argues that the application difficulties of the crime of infringing citizens' personal information under big-data conditions mainly arise from the changing structure of data collection, circulation, and use. Personal information is no longer limited to isolated identifiers but increasingly appears as aggregated, contextual, and reusable data. This change complicates the ability of courts to define personal information, assess seriousness, distinguish administrative illegality from criminal liability, allocate actor responsibility, and prove harm or causation.

The analysis shows that a single quantity-based approach is insufficient. Although quantity standards remain necessary for legal certainty, they cannot fully reflect the risks created by sensitive information, internal access abuse, platform outsourcing, technical circumvention, doxxing databases, or downstream fraud. The typical cases examined in this paper indicate that judicial reasoning should consider information type, data source, conduct method, actor role, illegal gains, and harmful consequences together. This comprehensive approach aligns more closely with the realities of big-data crime and better supports proportional punishment.

The paper therefore proposes several improvement paths. First, the identification of personal information should be based on identifiability and contextual risk while avoiding unlimited expansion of criminal-law protection. Second, the assessment of "serious circumstances" should combine quantity with sensitivity, source, method, and consequence. Third, the boundary between administrative violation and criminal offense should be clarified so that ordinary compliance defects are not over-criminalized while serious data abuse remains punishable. Fourth, evidence review should be strengthened, especially regarding data source, deduplication, identifiability, and causal connection. Fifth, corporate compliance and preventive governance should be improved through access control, log retention, outsourcing review, and internal accountability.

This study also has limitations. The analysis is mainly based on legal norms and selected typical cases, so it may not fully reflect all regional differences in judicial practice. Some case facts are limited by the public materials available. Future research may expand the sample of judgments, compare sentencing patterns across regions, and include administrative enforcement and non-prosecution cases. Overall, the crime should be applied in a way that protects personal information effectively while preserving the modest, precise, and supplementary role of criminal law.

References

1. Madhusudhanan, S., and Jose, A. C., "Privacy preservation techniques through data lifecycle: A comprehensive literature survey," *Computers & Security*, vol. 155, p. 104473, 2025.
2. Guo, Z., "Criminalisation of the illegal use of personal data: comparative approaches and the Chinese choice," *Humanities and Social Sciences Communications*, vol. 12, no. 1, pp. 1–16, 2025.
3. Bakhteev, D. V., Sosnovikova, A. M., and Kazenas, E. V., "Overcoming illegal cross-border transfer of personal data," *Journal of Digital Technologies and Law*, vol. 2, no. 4, pp. 943–972, 2024.
4. He, M., and Chen, Y., "Personal data protection in China: Progress, challenges and prospects in the age of big data and AI," *Telecommunications Policy*, p. 103076, 2025.
5. Gabel, M., Carrubba, C. J., Helmke, G., Martin, A. D., Staton, J. K., Ward, D., and Ziegler, J., "CompLaw: A Coding Protocol and Database for the Comparative Study of Judicial Review," *Journal of Law and Courts*, vol. 12, no. 2, pp. 466–492, 2024.
6. Oatley, G. C., "Themes in data mining, big data, and crime analytics," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 12, no. 2, p. e1432, 2022.
7. Brayne, S., "The criminal law and law enforcement implications of big data," *Annual Review of Law and Social Science*, vol. 14, no. 1, pp. 293–308, 2018.
8. Devins, C., Felin, T., Kauffman, S., and Koppl, R., "The law and big data," *Cornell JL & Public Policy*, vol. 27, p. 357, 2017.
9. Lei, C., "Legal control over big data criminal investigation," *Social Sciences in China*, vol. 40, no. 3, pp. 189–204, 2019.
10. Simmons, R., "Big data, machine judges, and the legitimacy of the criminal justice system," *UC Davis L. Rev.*, vol. 52, p. 1067, 2018.
11. Simmons, R., "Quantifying criminal procedure: how to unlock the potential of big data in our criminal justice system," *Mich. St. L. Rev.*, p. 947, 2016.
12. Henderson, S. E., "A few criminal justice big data rules," *Ohio St. J. Crim. L.*, vol. 15, p. 527, 2017.

13. Pramanik, M. I., Lau, R. Y., Yue, W. T., Ye, Y., and Li, C., "Big data analytics for security and criminal investigations," *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 7, no. 4, p. e1208, 2017.