Journal of Commercial Law Review
2025, VOL. 9, NO. 4, 11-14
DOI: 10.47297/wspjclrWSP2516-249702.20250904

Legal Challenges to Personal Information Security in the Big Data Era and AI Regulatory Approaches

Pengju Song

School of Economics and Management, Guangzhou Institute of Science and Technology, Guangzhou, Guangdong, 510000, China

ABSTRACT
In the era of big data, personal information security faces systemic challenges triggered by technological iteration, necessitating the reconstruction of a collaborative governance paradigm between law and technology. This paper reveals that personal information processing exhibits new characteristics: non-intrusive data harvesting, precise prediction, and large-scale abuse. Against this backdrop, the traditional legal framework, centred on "informed consent," encounters structural dilemmas such as subject abstraction, ambiguous liability, and regulatory lag. To address this, this paper proposes an AI-driven regulatory technology implementation pathway: establishing preemptive automated assessment mechanisms for data collection legitimacy using deep neural networks; deploying sensor networks for real-time anomaly detection in data streams under Federated Learning Frameworks; and introducing cross-chain innovative contract technology to enable cross-domain, chain-traversing provenance analysis and evidence collection for infringement activities. This system transforms legal rules into programmable regulatory nodes, forming a dynamic, closed-loop intelligent oversight mechanism with end-to-end traceability. It provides a technology-enabled institutional safeguard to balance the security of data flow with innovation development.

KEYWORDS
Big data; Personal information security; Legal challenges; Artificial intelligence; RegTech

1 Introduction

The deepening development of the digital economy has brought profound crises to the protection of personal information rights. The inherent demand for efficient data circulation drives collection technologies toward ubiquitous sensing—from imperceptible biometric capture in public spaces to continuous physiological monitoring via wearable devices—systematically dismantling individuals' physical boundaries and mental defences. Concurrently, machine learning-driven behavioural prediction models achieve molecular-level precision in mapping sensitive profiles, such as consumer tendencies and health risks, through multi-source data fusion, rendering individual autonomy nearly obsolete before algorithmic analysis. More gravely, the scaled application of decentralised storage topology and cloud computing has dramatically expanded the harm radius of data breaches. A single attack can trigger the fragmentation of privacy for hundreds of millions of users, while privacy-preserving computing-based sharing models have created new grey areas where data abuse becomes increasingly legalised. The existing legal framework exhibits structural inadequacy: informed consent mechanisms lose substantive efficacy under nested privacy policies and deceptive interaction patterns; the boundaries of responsibility among multiple stakeholders—platforms, data processors, and algorithm developers—become increasingly blurred through technological collaboration; and the iteration speed of legal rules lags behind the exponential growth of data processing technologies. While judicial rulings still determine jurisdiction based on the physical location of servers, edge computing nodes have already enabled the cross-border processing of data.

2 New Characteristics of Personal Information Security in the Big Data Era

2.1 The Imperceptible and Pervasive Nature of Information Collection

Modern information technology has established an environmental collection network that permeates all aspects of daily life. Surveillance systems in public areas continuously record pedestrian movement patterns using dynamic capture technology, while background programs on mobile devices automatically upload users' location and timing data. A typical manifestation is the covert operation of embedded sensor devices, where smart home appliances synchronously gather physiological changes in family members via temperature regulation modules, and in-vehicle systems generate driving behaviour models by analysing operational habits [1]. The essence of this technological implementation lies in bypassing manual interfaces, enabling the collection of information beyond users' conscious awareness. When wearable devices autonomously monitor physiological metrics during nighttime sleep, individuals lose immediate control over their own data. This pervasive environmental penetration renders information gathering an inherent property of physical spaces, compelling citizens to passively assume the role of digital donors against their will.

2.2 Precision User Profiling and Behavioural Prediction

Advanced data processing technologies are deconstructing the boundaries of human behavioural randomness. Commercial platforms precisely identify consumer decision-making sensitivity windows by analysing correlations between payment hesitation duration and page-switching frequency. Online education systems pinpoint knowledge gaps based on time spent on specific questions. The core capability lies in algorithmic decoding of behavioural patterns: combined modelling of login time fluctuations and contact list update frequency in social apps can project users' social relationship restructuring cycles; Healthcare platforms track department migration paths in appointment records to reflect health trend trajectories. Predictive mechanisms increasingly manifest as cross-domain parameter coupling: exercise frequency decay curves from fitness apps, combined with calorie intake increases from dining data, can trigger dynamic risk-level adjustments for financial products. As predictive models develop decision-making inertia through historical data iteration, individual choice freedom faces the real crisis of being swept along by algorithmically preset paths.

2.3 Scalable Risks of Data Leakage and Abuse

The fluidity of information elements enables security vulnerabilities to propagate through hierarchical structures. The topological characteristics of decentralised storage topology architectures mean a single security breach can trigger cross-border data leakage chain reactions—such as when e-commerce payment systems are compromised, leading to simultaneous leaks of transaction records and biometric data. The fundamental contradiction lies in the untraceability of secondary data utilisation: location data recorded by navigation software is resold through intermediary platforms, evolving into parameters for regional consumption capacity assessments; cross-analysis of community access control systems and delivery data generates profiles of solitary individuals for targeted marketing. Special risks stem from technological dependency chains. Vulnerabilities in generic functional modules (SDKs) can simultaneously expose thousands of applications to attack interfaces, creating ripple-effect risk propagation [2]. When encrypted communication technologies are used to circumvent legal oversight, identifying and blocking abusive practices often face technical failures. Security threats are now propagating from the personal information layer to the foundational trust of society.

3 Challenges Facing the Existing Personal Information Protection Legal Framework

3.1 Formalisation and Emptying of the "Informed Consent" Principle

The core contradiction in legal practice lies in the substantive failure of user authorisation mechanisms. Mobile applications often employ multi-layered, nested menu structures for critical permission clauses, requiring users to navigate through multiple interaction interfaces to understand the scope of consent fully. Interface interaction techniques are widely employed to disrupt decision-making: green visual cues guide users toward confirmation buttons for swift acceptance, while greyed-out cancel options are visually deemphasised. Pre-checked boxes are strategically embedded within dynamic visuals to divert attention. Procedural hollowing is further concealed within continuous modification mechanisms, where open-ended reservation clauses in service agreements grant operators unlimited rights to expand data usage—such as repurposing facial recognition permissions for commercial sentiment analysis. Systemic cognitive load barriers emerge as new obstacles: specialised privacy texts exceeding a hundred pages demand clause-by-clause confirmation, compelling most users to abandon substantive review. When checkbox actions devolve into mindless mechanical gestures, the foundational autonomy of legal decision-making collapses at the neurocognitive level.

3.2 Challenges in Identifying Responsible Parties and Burden of Proof

Technological synergy architectures create structural fractures in liability attribution. Smart device accidents often trigger a blame-shifting cycle among developers, data service providers, and hardware manufacturers. When navigation system errors cause traffic accidents, algorithmic logic and sensor accuracy become battlegrounds for mutual recrimination. This accountability dilemma stems from the black-box effect of technology: when pricing systems implement differential strategies, consumers cannot trace the origins of decision chains nor prove the price formation mechanisms at specific points in time [3]. There is a significant disconnect between evidentiary rules and their practical application. When users claim losses due to biometric errors, companies can often evade liability by merely presenting reports on the model's overall accuracy rate. A more profound contradiction stems from the non-personhood of technological entities. Iterative systems routinely delete underlying operational logs, leaving no accountable entity when tracing data processing actions from years prior. Within autonomous algorithmic systems, the concept of legal personhood is losing its corresponding physical embodiment.

3.3 The Conflict Between Lagging Legal Regulation and Rapid Technological Advancement

Legal regulatory frameworks face a perpetual catch-up dilemma. As regional data classification standards are being drafted, novel electromagnetic wave sensing technologies have already outpaced existing regulatory frameworks. Just as app store review standards are established, device-side independent processing solutions achieve localised data destruction architectures. The disconnect between judicial capabilities and technological advancement creates substantive barriers: privacy infringement scenarios enabled by edge computing prove difficult to capture with effective evidence chains. At the same time, traditional server logs can only extract obsolete instruction codes. The obsolescence cycle of standards continues to shorten. By the time regulators complete legislation restricting data collection for intelligent connected vehicles, steering wheel grip analysis technology will have already become a fundamental feature in the market. This speed disparity traps the legal system in a self-iterating paradoxical loop: by the time new regulations take effect, they often become outdated benchmarks for technological breakthroughs.

4 Application Pathways of Artificial Intelligence in Personal Information Security Regulation

4.1 Leveraging AI Technology for Automated Preemptive Risk Review

Deep learning-driven compliance review systems are emerging as the technological vehicles for legal implementation. Neural network-based semantic analysis models can automatically deconstruct ambiguous clauses in privacy agreements, triggering risk flags for open-ended statements such as "optimising user experience" or "enhancing service quality." By constructing enterprise data behaviour knowledge graphs, the system continuously refines predictive capabilities. It compares the declared scope of temperature collection functions in smart home devices against actual transmitted data dimensions. When undeclared parameters, such as EEG monitoring, are identified in transmission packets, the device certificate automatically enters a temporary freeze state. The core innovation lies in implementing an algorithmic ethical pre-screening framework. During the development of the user profiling system, adversarial training is used to establish discrimination factor detection models. By employing perturbation techniques to disrupt hidden correlations between gender labels and credit limit allocation parameters, the system gains ethical compliance assurance before deployment [4]. This algorithmic compliance intervention automatically aligns the scale of data processing with statutory approval tiers, transforming the programmable compliance system into a technical barrier against systemic violations. The preemptive technical firewall thus bridges the gap between legal norms and technical implementation, establishing a technological foundation for institutional safeguards at the outset of the data lifecycle.

4.2 Building an AI-Based Real-Time Data Flow Anomaly Monitoring System

A distributed collaborative regulatory architecture reshapes the visibility of data flow processes. Within a dual-layer network comprising device-side validation modules and cloud-based intelligent hubs, terminal sensors continuously verify whether data processing aligns with declared scenarios. For instance, when a smart elderly wristwatch transmits indoor positioning data beyond its health monitoring function, geofencing protection protocols immediately trigger local data blocking strategies. The central system employs a federated analytics architecture for encrypted monitoring of multi-source data. When hospitals conduct large-scale access to terminal patient medication records for quantitative analysis models, anomaly detection engines trace back to the source of demand and activate data circuit-breaker mechanisms for non-research institution accounts. Breakthrough progress is evident in the visual representation of the regulatory chain: Companion digital tokens are generated during the flow of student classroom behaviour data. When educational institutions access attention concentration metrics, terminals mandatorily generate cryptographic evidence containing usage duration and purpose, simultaneously sending data flow confirmation alerts to parental accounts. This tamper-proof audit, accompanied by a tamper-proof audit, achieves legally mandated transparent management loops, resolving traditional oversight delays while preventing human interference with technological innovation. A bulletproof-glass-like transparent regulatory environment makes "data usage as evidence" and "application as trace" standard operational norms.

4.3 Empowering AI Technology for Post-Incident Infringement Tracing and Forensics

Artificial intelligence facilitates the collection of infringement evidence, transitioning from superficial tracing to root-cause analysis. In investigating cross-border illegal data transfer cases, the spatio-temporal topology analysis engine constructs multidimensional mapping models to identify anomalies in encrypted tunnel protocols: instances of encryption standard downgrading during data transmission are recorded as signature technical traces of illegal transfers. Protocol stack deconstruction pinpoints the originating node of operational commands. Breakthroughs in biometric

forgery detection centre on modelling physical law. Anomalous periodic changes in iris texture within deepfake videos are validated via fractal algorithms as violating light field diffraction laws, thereby forming a legally admissible chain of evidence for image tampering. Revolutionary applications manifest in programmable evidence systems: When smart car manufacturers store passenger conversations under security pretexts, vehicle control signal decoders can prove that door lock statuses and transmission engagement modes during specific periods failed to meet warning trigger conditions, thereby directly forming specialised evidence packages that prove false demand claims [5]. This three-dimensional testimony, integrating physical attributes with digital imprints, enables legal accountability to transcend the limitations of technological black boxes. When data recipients are confronted with three-dimensional simulation models that reconstruct transmission paths during judicial proceedings, infringers' defences often collapse entirely before the technical evidence system. The digital encapsulation of objective illegal elements provides a supportive framework for legal sanctions.

5 Conclusion

The deep integration of artificial intelligence into personal information security governance marks a fundamental shift in regulatory models from reactive remediation to proactive defence. By establishing a technological oversight chain spanning the entire data lifecycle—embedding programmable compliance pre-screening mechanisms during collection, deploying distributed, persistent verification systems during circulation, and implementing multidimensional spatiotemporal evidence anchoring upon infringement—the systematic coupling of technical rules and legal norms forms a dynamic governance barrier. This technology-embedded governance paradigm safeguards individuals' right to self-determination over their personal information while laying the institutional groundwork for digital economic development that strikes a balance between security and innovation. Ultimately, it achieves a complete closed-loop transformation of personal information protection—from abstract legislative intent to concrete technological implementation.

About the Author

Pengju Song, PhD, Distinguished Associate Professor. Research Interests: Judicial Systems, Legislation and Corporate Governance.

References

[1] Digital Wellbeing Safeguards: Protecting Personal Information Security in Consumer Sectors Yields Tangible Results [J]. China Cyberspace News, 2025, (07): 34-35.

[2] Zhang Yuhang. Conflict and Regulation Between Big Data Investigation and Personal Information Protection [J]. Journal of Changzhou University (Social Sciences Edition), 2025, 26(01): 43-54.

[3] Mao Yixiao. Full-Process Specialised Control Mechanism for Personal Information Security Risks [J]. Jiangsu Social Sciences, 2024, (06): 160-168.

[4] Yang Fei. Concealment and Disclosure: On Algorithms' Impact on Personal Information Security and Countermeasures—An Exploration from the Perspective of Media Visibility [J]. Journal of Intelligence, 2024, 43(10): 157-165.

[5] Xu Wei, Li Wenmin. Security Risks and Legal Responses to Personal Information Collection by Apps in the Digital Intelligence Era [J]. Journal of Shandong Administrative College, 2025, (01): 113-121.