

THE PARADIGM SHIFT AND THEORETICAL ADAPTATION OF JUDICIAL JUSTICE EVALUATION UNDER DIGITAL AND INTELLIGENT EMPOWERMENT

Zhuang Xinyu *

I. INTRODUCTION

The *Recommendations of the Central Committee of the Communist Party of China for Formulating the 15th Five-Year Plan for National Economic and Social Development*, adopted at the Fourth Plenary Session of the 20th Central Committee of the Communist Party of China, state that ‘We should better regulate the exercise of judicial power, refine the systems through which judicial justice is delivered and evaluated, and do more to ensure the fairness, consistency, and authority of judicial decisions.’¹ As a systematic examination of the effectiveness of judicial reform, the regulated exercise of judicial power, and the quality of judicial cases,² judicial justice evaluation refers to ‘the sum total of activities through which diverse actors in society judge, perceive, and assess the justice of judicial organs and their judicial activities’.³ At present, under digital and intelligent empowerment, China’s judicial business systems are undergoing iterative upgrading from traditional informatization to data-based and intelligent operation. Judicial justice evaluation has accordingly undergone adjustments in its methods of data collection, feedback mechanisms, and foundations of credibility, and its governance paradigm is shifting toward a data-driven model of governance. Therefore, existing theories of judicial justice evaluation should undergo corresponding interpretive renewal and conceptual reconstruction to provide theoretical support for building an independent knowledge system of judicial justice evaluation.

* Ph.D. Candidate at the School of Criminal Justice, China University of Political Science and Law. This paper is supported by the Fundamental Research Funds for the Central Universities, entitled *Research on the New Criminology Theory System from the Perspective of Chinese Modernization of Crime Control* (24CXTD02) and the General Project for Procuratorial Theory Research of the Supreme People’s Procuratorate in 2022, entitled *Research on the Evaluation Indicator System for Procuratorial Work* (GJ2022C33).

1 *Recommendations of the Central Committee of the Communist Party of China for Formulating the 15th Five-Year Plan for National Economic and Social Development*, People’s Daily, October 29, 2025, at 1.

2 See Jiang Yinhua, *The Thinking Turn in Judicial Evaluation Studies*, 23(2) Journal of National Prosecutors College 3-11 (2015).

3 See Chen Guangzhong, *Building a Higher Level of Judicial Justice*, 1 China Legal Science 20 (2026).

II. THE UNDERLYING LOGIC OF JUDICIAL JUSTICE EVALUATION UNDER DIGITAL AND INTELLIGENT EMPOWERMENT

Under digital and intelligent empowerment, the operational logic of judicial justice evaluation lies in the data-enabled reconfiguration of its value foundation, normative basis, and factual support. To clarify this logic, it is necessary to explain both why the justice of judicial activity is amenable to evaluation and how digital and intelligent technologies can participate in and substantively shape that evaluation.

A. Why Judicial Justice Is Amenable to Evaluation

Evaluation is a cognitive activity involving value judgment and guidance. Judicial justice is amenable to evaluation because judicial decisions are produced within an established normative framework, while the operation of judicial procedures and reasoning of judicial decisions can be fully presented through channels of judicial openness and subjected to scrutiny. On this basis, judicial justice evaluation is supported at the levels of value, normativity, and fact.

First, at the value level, the justice of judicial decisions establishes an evaluative benchmark. The judiciary performs the important function of safeguarding social fairness and justice, and judicial justice evaluation accordingly takes justice as its core criterion while, under the guidance of that value, also taking into account considerations such as efficiency, integrity, and people-centeredness. This corresponds to the legitimate expectations of the public with respect to the judiciary.⁴

Second, at the normative level, the consistency and norm conformity of judicial decisions provide a normative benchmark for evaluation. Normative systems, including intra-Party regulations, organizational rules governing judicial organs, and norms of judicial professional ethics, continuously define the proper boundaries, procedural requirements, and allocation of responsibility in the exercise of judicial power, thereby making it possible to compare, measure, and test judicial decisions within a common normative framework.

Third, at the factual level, the traceability and openness of judicial

⁴ See Wu Yingzi, *On Judicial Recognition: Crisis and Reconstruction*, 3 China Legal Science 186-206 (2016).

materials provide observable carriers for evaluation. Judicial outcomes are realized through litigation procedures and are fixed and preserved in the form of case files, judicial documents, and judicial recommendations. The integrity and traceability of case files, the openness and verifiability of judicial documents, and the exemplary and governance-oriented nature of judicial recommendations make it possible to present and assess judicial effects in terms of the chain of fact-finding, logic of evidentiary admission, structure of legal application, and feedback from social governance.

In light of the differences in scope and functional orientation, judicial justice evaluation may be categorized into overall evaluation, management evaluation, and specialized evaluation.⁵

Overall evaluation is a comprehensive assessment that takes the judicial system and the process of judicial reform as its object, with a primary concern for the macro-level examination of judicial power. It is typically organized and conducted by authoritative institutions or research institutes and, through index-based measurement, annual reports, or thematic assessments, presents trends over a given period in the effectiveness of judicial reform, the state of institutional operation, and the overall level of the rule of law.

Management evaluation is an assessment oriented toward internal governance and responsibility implementation, taking the organizational operation of judicial organs and the performance of judicial personnel as its object.⁶ Management evaluation encompasses both judicial performance assessment in the administrative sense and judicial quality-and-effectiveness evaluation in the case-management sense, including case-quality assessment. Through the continuous monitoring of elements such as the fulfillment of duties, quality and efficiency of case handling, and compliance with norms, it seeks to optimize the management of judicial operations and strengthen accountability constraints.

Specialized evaluation is an application-oriented assessment that focuses on specific policy fields or governance issues and relies primarily on adjudicative data and case-handling information. In a broader governance

5 See GAO ZHIGANG, RESEARCH ON THE EVALUATION SYSTEM OF THE JUDICIAL SYSTEM, at 53-84 (Shanghai People's Publishing House, 2018).

6 For purposes of this article, 'judicial organs' is used in a relatively broad sense consistent with China's state institutional design, referring to the people's courts and the people's procuratorates, but not to public security organs.

sense, such specialized evaluations may also be reflected in assessments that use judicial data or dispute resolution indicators to examine governance performance in particular fields. Focusing on governance tasks concerning social risks in a specific field, it examines the application and operational effects of judicial rules and emphasizes targeted diagnosis, comparative analysis, and improvement-oriented conclusions regarding particular problems.

Accordingly, these three forms of evaluation correspond to different organizational actors and positions within governance. Overall evaluation is often advanced by the public, legal professional community, and external third-party organizations. Management evaluation is mainly conducted by internal business management departments within judicial organs. In contrast, specialized evaluation is commonly organized by domestic or international professional institutions or interdepartmental mechanisms in response to governance needs in particular fields. Under this structure of functional differentiation and multi-actor evaluation, judicial justice evaluation can produce mutually corroborative conclusions and can also transform those conclusions into argumentative grounds for institutional improvement in a verifiable manner.

B. How Digital and Intelligent Technologies Empower Judicial Justice Evaluation

General Secretary *Xi Jinping* has stressed the need to ‘promote wide application of big data, artificial intelligence, and other new technologies in our judicial work’.⁷ Currently, digital and intelligent technologies have become important for the operation of judicial institutions and for updating the paradigm of judicial justice evaluation. Digital and intelligent empowerment is scientifically grounded and practically feasible because it transforms unstructured judicial information into structured, evaluable data materials, thereby enabling overall evaluation, management evaluation, and specialized evaluation to generate credible evaluative conclusions on a verifiable basis.⁸

In this article, ‘digital and intelligent empowerment’ refers to a set

7 XI JINPING, SAFEGUARD POLITICAL SECURITY, SOCIAL ORDER AND PEACEFUL LIVES (JANUARY 15, 2019), in XI JINPING, THE GOVERNANCE OF CHINA III, at 411 (Foreign Languages Press, 2020).

8 See M. Simmler, G. Canova & K. Schedler, *Smart Criminal Justice: Phenomena and Normative Requirements*, 2 International Review of Administrative Sciences 422-423 (2023).

of technological gains achieved through digital platforms that aggregate judicial data and business applications and that, with the support of big data, blockchain, and artificial intelligence, enable process logging, retrieval and comparison, and data mining, so that rule requirements, case-handling conduct, and output results are traceable and verifiable.⁹ In terms of its specific support for judicial justice evaluation, such empowerment may be summarized into three categories.

First, service-and-aggregation empowerment. Electronic litigation service platforms enable one-stop online handling of a wide range of litigation matters, including filing, mediation, preservation measures, and enforcement, as well as cross-regional litigation matters, thereby translating judicial objectives, such as accessibility for the public and transparency, into concrete services, including online hearings and intelligent support for people's assessors.¹⁰ The data generated through these services provide experiential evidence and factual support for the overall evaluation of judicial operations.

Second, recording-and-computation empowerment. Against the backdrop of the Supreme People's Court's efforts to build a nationwide 'single network' for courts, electronic litigation service platforms in different localities have become interconnected and store full-process judicial data. The procuratorial organs' Unified Business Application System 2.0, through functions such as intelligent case-card completion,¹¹ case-review push notifications, and alerts for procedural time-limit overruns, together with process monitoring, quality assessment, and prosecutorial supervision, integrates full-process recording and computational analysis across procuratorial work. Thus, it provides a verifiable data foundation and a more precise performance profile for management evaluation.¹²

Third, knowledge-supply empowerment. Public case databases, such as the database of adjudication documents and case repositories of the people's courts, as

9 See Wang Bing, *What Is 'Digital Intelligence'? A Study of the Multiple Meanings of the Concept of Digital Intelligence*, 7 *Journal of Intelligence* 75-76 (2023).

10 See Gong Shanyao, *The Modernization and Future Orientation of Chinese Judicial Democracy in the Age of Digital Intelligence*, 5 *Legal Forum* 131-140 (2024).

11 See Shen Guojun, Li Yi & Wang Yongzheng et al., *Paths for Optimizing Case Cards in Procuratorial Organs*, 24 *People's Procuratorial Semimonthly* 13 (2024).

12 See Shi Zhaokun, *Procuratorial Business Application System 2.0: Letting Machines Do 'Addition' and Humans Do 'Subtraction'*, *Procuratorial Daily*, November 4, 2021, at 2.

well as internal guidance platforms, including the FaDa and JianDa platforms,¹³ continuously provide clarified conclusions regarding the rules of similar cases and key difficulties in application through the dissemination of similar case guidance and centralized Questions and Answers (Q&A). This provides a shared and comparable scale for adjudicative reasoning and case-handling essentials and offers reliable materials for specialized evaluation.

The three forms of empowerment do not affect all types of judicial justice evaluation in the same way. Service-and-aggregation empowerment is more readily directed to overall evaluation because online litigation services, public-access inquiry mechanisms, and convenience-oriented arrangements can translate requirements such as service accessibility and procedural openness into continuously logged data. Recording-and-computation empowerment is more closely aligned with management evaluation. By relying on the automatic logging of business nodes, the synchronized aggregation of procedural data, and the automatic verification of rule-based conditions, it enables evaluation in the assessment of judicial quality and effectiveness to move beyond result-based calculation toward process-based review. Knowledge-supply empowerment, in turn, provides specialized evaluation with publicly accessible data sources. Through similar case recommendations, knowledge graphs, business-oriented Q&A, and rule prompts, it offers more concentrated points of reference for judgment in relation to particular fields, matters, and risk points. Accordingly, what digital and intelligent empowerment changes is not merely the quantity of evaluative materials or the speed of their processing, but more importantly, the fact that it provides distinct types of judicial justice evaluation at different levels and for different objects, with corresponding data foundations and knowledge support.

III. THE PARADIGM SHIFT TOWARD DATA-DRIVEN GOVERNANCE IN JUDICIAL JUSTICE EVALUATION UNDER DIGITAL AND INTELLIGENT EMPOWERMENT

Under digital and intelligent conditions, judicial business platforms continuously convert information on reforms in judicial operations, information on the performance of case-handling duties, and adjudicative information in specific fields into evaluative data. On this basis, judicial

13 The FaDa and JianDa platforms are the internal Q&A and guidance platforms of the people's courts and people's procuratorates.

justice evaluation under digital and intelligent empowerment has undergone a paradigm shift toward data-driven governance. The paradigm of data-driven governance does not merely refer to a greater degree of reliance on data in judicial justice evaluation; rather, it signifies a transformation in the logic through which evaluation is generated. Under the traditional model, evaluation is based primarily on *ex post* aggregation, manual reporting, and periodic statistics, with data serving chiefly as a vehicle for presenting results. However, under conditions of digital and intelligent empowerment, information on procedural nodes, procedural actions, document contents, and managerial traces generated during judicial operation can be synchronously retained within a unified business system. The materials on which the evaluation relies thus no longer derive mainly from *ex post* compilation but increasingly from the case-handling process itself. As a result, judicial justice evaluation is gradually forming a new operational chain grounded in process logging, conditioned on the embedding of rules, and characterized by identification, comparison, feedback, and correction, thereby progressively shifting from a result-aggregation model to one centered on process logging, rule embedding, and feedback loops.

A. Transformation in Data Collection:

From Manual Extraction to Human-Machine Collaboration

Judicial justice evaluation is primarily manifested in the qualitative categorization and quantitative measurement of judicial business data. Its core method of data collection lies in the accurate and efficient collection, cleansing, and analysis of judicial business data. Traditional judicial justice evaluation has largely relied on manually recorded data, aggregated reporting forms, and sampling-based verification. However, dispersed counting rules, data omissions, and delayed feedback often undermine the robustness of subsequent analysis.

With the advent of digital and intelligent empowerment, the data required for evaluation are increasingly embedded in the routine operation of business systems. Such data are automatically generated, logged, and accumulated as continuous records; accordingly, data acquisition shifts from *ex post* aggregation to process-based recording. Human-machine collaboration in judicial justice evaluation is mainly reflected in the support provided by digital and intelligent technologies in data extraction and structuring. Judicial business

systems may employ Optical Character Recognition (OCR) to recognize text and Natural Language Processing (NLP) to extract case elements, and further utilize deep-learning models such as Bidirectional Encoder Representations from Transformers (BERT) and Enhanced Representation through Knowledge Integration (ERNIE) to conduct associative analysis between statutory provisions and cases; together with knowledge graphs, these technologies facilitate retrieval and comparison, relational linking, and anomaly alerts in the collection, storage, and processing of data.¹⁴

Take the performance-evaluation module of the Procuratorial Business Application System 2.0 as an example. In addition to automatically imported data, the system includes supplementary procedures, such as self-reporting of workload. This is intended to connect system logs with individualized explanations of contributions, thereby preserving records of counting rules while avoiding the automatic equating of system-generated logs with the full picture of performance. It also leaves the necessary narrative and explanatory space for complex performance scenarios.

Under human-machine collaboration conditions, evaluative materials are increasingly generated directly from business processes. Information concerning case acceptance, procedural progression, document preparation, supervisory correction, and feedback on outcomes can be continuously accumulated within the system. Therefore, evaluators confront not merely *ex post* numerical results compiled after the fact but process-based materials that can be traced back to specific procedural nodes, documents, and operational traces. Correspondingly, the object of judicial justice evaluation expands from a static comparison of outcomes to the continuous observation of whether procedural operation is compliant, performance-related actions are adequate, and managerial requirements have been implemented. In other words, what human-machine collaboration brings about is not merely greater

14 OCR refers to a data-processing technology that converts textual information in images and scanned documents into editable and searchable text through image processing and pattern recognition. NLP refers to a technological system that uses computer techniques to segment, understand, extract, analyze, and generate human language. BERT is a bidirectional pre-trained language model based on the transformer architecture, capable of improving performance in tasks such as text understanding, classification, and matching by incorporating contextual information. ERNIE is a pre-trained language model that incorporates knowledge-enhancement mechanisms and emphasizes the improvement of Chinese language understanding through the integration of lexical, entity-based, and semantic knowledge. A knowledge graph is a structured knowledge-representation system constructed on the basis of entities, attributes, and relations, which can present semantic associations among objects in graph form and support retrieval, inference, and relational analysis. See Li Xueyao, *Judicial Bias and Cognitive Intervention in the Application of Large Language Models*, 5 Political Science and Law 65-76 (2025).

efficiency in data extraction, but a transformation whereby judicial justice evaluation shifts from the human production of evaluative materials to the direct generation of such materials by the case-handling process itself.

*B. Transformation in Data Feedback:
From Responsive Evaluation to Agile Evaluation*

Judicial justice evaluation renders value judgments on judicial operations against specified standards and serves functions such as demonstrating reform achievements, strengthening managerial supervision, and promoting improvements in case-handling quality. Traditional responsive evaluation is often triggered by external concerns or salient problems; information is periodically aggregated, assessed in a centralized manner, and then fed back as directions for improvement. While this facilitates the production of phased profiles, it typically relies on *ex post* review, creating a time lag between evaluation and operation.¹⁵

Digital and intelligent empowerment reshapes both the mode of generating evaluative information and the temporal structure of the evaluation. Under conditions of human-machine collaboration, activities of adjudication, management, and service are retained as data within business systems; evaluative materials are generated synchronously with judicial operations, and periodic aggregation is no longer the sole precondition for initiating evaluation. Accordingly, judicial justice evaluation shifts from an emphasis on *ex post* response to an agile mode that stresses process tracking, timely identification, and dynamic assessment.

Agile evaluation may be understood as the application of the idea of agile governance to judicial justice evaluation: relying on an online and same-source data foundation, embedding evaluation into the daily operational chain through cross-departmental and cross-level coordination, continuously observing key links, and promptly identifying and addressing abnormal deviations, thereby improving responsiveness and correction efficiency.¹⁶

The long-standing predominance of responsive evaluation is rooted in

15 See Cui Zhiwei, *The 'Responsive' Turn in Criminal Justice: Seeking the Basis for the Substantive Reasonableness of Punishment*, 2 Hebei Law Science 166-177 (2019).

16 See Zeng Runxi & Zhu Xingtong, *Research on the Path of Cyberspace Governance Based on the Concept of Agile Governance*, 1 Journal of Hunan University (Social Sciences) 147-154 (2026).

the fact that the information available to traditional judicial administration is lagging and periodic in nature. Whether in the form of special inspections, quarterly circulars, or year-end assessments, an overall judgment typically cannot be formed until the end of a given period when reports, aggregated materials, and sampled case reviews become available. Thus, to a considerable extent, evaluation serves the functions of *ex post* summarization and pressure transmission. Under such conditions, evaluation may identify problems, but it is difficult to intervene in a timely manner during the formation of those problems.

In contrast, agile evaluation is based on the real-time logging of business operations, continuous aggregation of data, and automatic triggering of rules. It does not need to wait until the completion of an entire cycle before making a unified assessment; rather, it can generate earlier and faster feedback with respect to abnormal fluctuations, emerging risks, and operational states that deviate from applicable rules. Accordingly, the temporal point of judicial justice evaluation shifts from periodic and summative nodes to the very process of judicial operation itself, and its function correspondingly moves beyond a mere response to existing problems toward a continuous governance tool that also serves warning, correction, and guidance functions.

Within an agile evaluation framework, functional emphases across different evaluation targets become more differentiated and are supported by richer materials. First, in evaluating reform achievements and the state of institutional operation, online operations and same-source data enable continuous presentation of operational trends and structural features. Thus, evaluation moves from periodic stock-taking to rolling observation and trend explanation, significantly enhancing its capacity to account for the state of judicial operations.

Second, when evaluating the performance quality of judicial organs and their personnel, process logging aligns duty lists with conduct at specific procedural nodes, enabling the evaluation to focus more directly on process verification and the confirmation of responsibility-related facts rather than merely on outcome-based statistics.¹⁷

Third, in case-quality assessment, business guidance platforms and similar case recommendation platforms provide normative references, whereas case files, judicial documents, and process logs provide reviewable materials. As

17 See Gao Jingfeng, *Practical Paths for Deepening the Determination and Pursuit of Judicial Responsibility in Procuratorial Organs*, 1 *Journal of the National Prosecutors College* 117-130 (2025).

these two types of materials are consolidated within the same system, evaluation can identify weaknesses in case quality at an earlier stage, promptly generate alerts and review leads, and move quality control forward from terminal relief mechanisms to the case-handling process itself, while ensuring that quality judgments rest on reasons and materials that are checkable and reviewable.

C. Transformation in Data Credibility: From Measurability to Verifiability

In the context of judicial justice evaluation, data credibility may be summarized as the pursuit of three forms of trust. Subject-based trust emphasizes that data can be mapped onto specific node-level conduct and documentary materials, facilitating checking and re-examination. Institutional trust emphasizes that extraction boundaries, counting rules, and logging standards are clear and stable. Organizational trust emphasizes that data can be accessed across levels and departments in a manner that supports consistent judgments.¹⁸

Influenced by New Public Management, judicial justice evaluation has long relied on measurable indicators to present reform achievements and the state of institutional operation, gauge the performance of organs and personnel, and provide an overall assessment of case handling.¹⁹ Technological advances brought by digital and intelligent empowerment have reshaped the generation and preservation of data: indicators, their generation rules, data sources, and process records become aligned on a same-source basis. Thus, evaluative materials shift from outcome data aggregated after the fact to process records supported by logged workflows and retrievable data snapshots.

Accordingly, measurability is further extended to verifiability. That is, evaluation not only presents changes in indicators but can also trace those changes back to specific procedural steps, judicial operations, and documentary materials, and subject them to review under unified counting rules. This provides shared support for realizing subject-based trust, institutional trust, and organizational trust in relation to the relevant data. Under the requirement of handling every case with high quality and efficiency,

18 See Qi Zhiwei, *The 'Trustworthy' Data Circulation Mechanism of Digital Government and Its Realization*, 5 Journal of Beijing Administration Institute 95-104 (2025).

19 See Zhu Qianwei, *The Governance Approach of Post-New Public Management Theory: Theory and Reality*, 2 Fudan Public Administration Review 251-268 (2022).

judicial quality-and-effectiveness evaluation, as an evaluative practice developed in the Chinese judicial context, has formed an evaluative chain in which indicators serve as clues, process materials function as support, and conclusions are supported through reason-giving.²⁰ It also interfaces with overall evaluation, management evaluation, and specialized evaluation.

Verifiability better captures the quality requirements of judicial justice evaluation under conditions of digital and intelligent empowerment because the credibility of such evaluation does not depend on whether a score is numerically precise but on whether the conclusion can be questioned, explained, and reviewed. In terms of subject-based trust, evaluative conclusions should not remain at the level of abstract rankings or score labels; they should be traceable to specific cases, procedural nodes, and material content so that relevant actors may understand how the conclusion was reached. In terms of institutional trust, evaluative standards should not only be capable of repeated application, but also be able to explain their conditions of application, statistical criteria, and methods for handling exceptions to prevent the same indicator from shifting in meaning across different contexts. In terms of organizational trust, superior and subordinate organs, different operational lines, and relevant supervisory bodies should be able to form mutually intelligible judgments based on the same evidentiary materials and rule logic, rather than offering fragmented interpretations based solely on aggregated reports. Accordingly, under digital and intelligent empowerment, the quality standard of judicial justice evaluation is shifting from whether it can be calculated to whether it can be traced, adequately explained, and effectively reviewed.

IV. THE THEORETICAL CHALLENGES OF DATA-DRIVEN JUDICIAL JUSTICE EVALUATION

As discussed above, the paradigm shift toward data-driven governance in judicial justice evaluation under digital and intelligent empowerment has structurally reshaped the generation of evaluative materials, feedback mechanisms, and foundations of credibility, thereby expanding the scope of evaluation and improving operational efficiency. Yet paradigm transformation yields more than technical gains; it also gives rise to theoretical challenges,

20 See Zhang Jie, *On the Systematic Evaluation of Judicial Quality and Effectiveness*, 8 Social Sciences 161-171 (2025).

including difficulties in constructing an explanatory framework, an imbalance in the evaluative value structure and the need for further refinement of proof rules.

*A. The Explanatory Framework Challenge:
Limited Visibility of the Objective Rules of Judicial Activities in
Quantitative Indicators*

General Secretary *Xi Jinping* has emphasized that ‘When improving and reforming our judiciary, we must follow the objective rules of judicial activities.’²¹ Digital and intelligent empowerment has substantially reduced the costs of collecting and computationally analyzing judicial business data, making the organization of evaluative materials more efficient. Nevertheless, the explanatory foundation of evaluation must still accord with the objective rules of judicial activities. The objective rules of judicial activities, as used in this article, refer to rules that objectively exist and repeatedly occur in the judicial process, determine the trajectory of judicial activity and the quality of adjudication, and must be strictly observed.²² They perform two principal functions in judicial justice evaluation. First, they set normative boundaries for the determination of evaluative goals, selection of indicators, and specification of counting rules. Second, they provide an indispensable procedural context for explaining differences and attributing responsibility.

However, the rules governing judicial activities are typically embedded in the case-handling process through procedural design, judicial reasoning, and responsibility allocation; therefore, they are difficult to translate directly and equivalently into numerical indicators.²³ Under digital and intelligent empowerment, indicatorization often compresses judicial phenomena, laden with normative conditions and procedural context, into data symbols. Such compression enhances comparability, but it can also diminish the visibility of those rules. Fluctuations in a single indicator may correspond to multiple procedural pathways and case conditions, while aggregation may smooth over differences in factual complexity, procedural choices, and the protection of parties’

21 XI JINPING, REFORM THE JUDICIAL SYSTEM (MARCH 24, 2015), in XI JINPING, THE GOVERNANCE OF CHINA II, at 142 (Foreign Languages Press, 2018).

22 See Xiang Zexuan & Tan Qingzhi, *The Objective Rules of Judicial Activities and Procuratorial Reform*, 5 Tribune of Political Science and Law 107-118 (2009).

23 See Ji Weidong, *China’s Judicial Modernization Driven by Data and Algorithms*, 4 China Law Review 87-101 (2024).

rights, loosening the linkage between data outputs and normative evaluation.

Accordingly, the rules governing judicial activities cannot be translated with equal ease into evaluative indicators at every level. With respect to matters that are more strongly procedural in nature and have relatively clear boundaries, such as case-handling time limits, procedural nodes, service of documents, and file archiving, indicator-based treatment is usually comparatively straightforward, because the object of assessment is relatively definite and the governing conditions are comparatively stable. By contrast, with respect to matters such as whether procedural discretion has been exercised appropriately, whether legal reasoning and explanation have been sufficient, whether evidentiary review has been thorough, and whether rights protection has been adequately secured, the rules governing judicial activities often require multifactorial judgment. Such judgments depend on an integrated assessment of the specific facts of the case, the procedural context, and the application of law, and they are therefore difficult to compress into a uniform standard of valuation. This is especially true in difficult and complex cases, where judicial justice is often reflected precisely in the prudent handling of exceptional circumstances, the balanced weighing of procedural interests, and the adequate response to case-specific differences. If one were nevertheless to continue following the logic of linear decomposition, uniform measurement, and horizontal comparison favored by New Public Management, matters that ought to be assessed through legal reasoning would all too easily be miscast as managerial matters capable of direct counting and simple comparison.

As large judicial models enter adjudicative assistance and managerial analysis, model inferences rely even more on structured inputs and preset labels; contextual differences may therefore be translated into ambiguous outputs, thereby weakening the justificatory force and replicability of evaluative conclusions. New Public Management is unable to fully explain judicial justice evaluation under conditions of digital and intelligent empowerment precisely because, although it is well-suited to matters such as goal decomposition, process control, and performance comparison, it is far less capable of explaining how those aspects of judicial activity that depend on legal reasoning, procedural safeguards, and case-specific judgment should enter the evaluative framework. Digital and intelligent technologies undoubtedly enhance capacities for data extraction, information comparison, and risk identification, but this does not mean that the rules governing judicial activities have thereby been fully

revealed. On the contrary, the stronger the data-processing capacity becomes, the more necessary it is, at the theoretical level, to determine which judgments may properly be represented through indicators and which must remain within the domain of legal interpretation and human review. Otherwise, the more sophisticated the evaluative system becomes, the greater the risk that departures from the rules governing judicial activities may be concealed.

*B. The Value Structure Challenge:
The Need to Further Incorporate Digital Justice into
Judicial Professional Ethics*

Under the influence of managerial judicial reform, judicial justice evaluation often quantifies judicial activity through efficiency-oriented indicators and allocates resources and responsibility accordingly.²⁴ Yet the value structure of judicial justice evaluation cannot be reduced to efficiency alone. It should also coordinate multiple values such as quality, rationality, and integrity within the overarching framework of judicial justice.²⁵ Digital and intelligent empowerment makes it easier for judicial values to be embedded and stabilized in computable rule structures. Business systems convert procedural nodes, case-handling time limits, document generation, and service matters into continuous records. Because efficiency-related variables are easier to extract, measure, and compare, they are more likely to dominate indicator systems. By contrast, elements such as procedural fairness, the quality of judicial reasoning, and the accessibility of remedies are difficult to quantify directly and are often reflected only through proxy indicators, sampling-based review, or *ex post* verification.

As a result, judicial justice evaluation may gradually gravitate toward efficiency variables and, through repeated computation, solidify an efficiency-centered evaluative orientation. Under such circumstances, the value orientation underlying evaluation is embedded in system operation through indicator rules and model logic, and ultimately projected via professional ethics into judicial personnel's everyday choices in performing official duties. Existing

24 See Guo Song, *The Realistic Landscape and Holistic Regulation of Managerial Judicial Reform in China*, 6 *Tsinghua University Law Journal* 109-126 (2024).

25 See Xiao Yi, *On the Functional Positioning of Value Judgment in Judicial Adjudication*, 1 *Law and Social Development* 124-145 (2026).

written norms of judicial professional ethics focus primarily on traditional requirements such as loyalty, justice, integrity, and people-centeredness, are supplemented by behavioral guidance on civility, and have yet to fully incorporate the digital ethics requirements for agile evaluation.²⁶

Although the field of digital governance has already developed principled requirements aimed at avoiding algorithmic bias, mitigating black-box effects, and preventing governance overreach,²⁷ these requirements have not yet been systematically incorporated into norms of judicial professional ethics and thus cannot be translated into clearly applicable professional duties and clearly delimited boundaries of responsibility. Consequently, the institutional advantage of efficiency variables is more readily reinforced by digital and intelligent evaluative mechanisms, whereas value elements associated with digital justice, such as transparency and explainability, reviewability of rules, and correctability of deviations, lack corresponding institutional support and therefore struggle to impose constraints of comparable strength on indicator design, weight allocation, and the use of results.

*C. The Proof Rules Challenge:
Weak Mechanisms for Objections and Responsibility in
Digitally Enabled Evaluation*

The challenge concerning proof rules in judicial justice evaluation under digital and intelligent empowerment is reflected in the difficulty of translating data credibility into the credibility of evaluative conclusions. Through same-source logging, clarified counting rules, and shared access, digital and intelligent empowerment enhances the checkability and reviewability of data, thereby supporting subject-based trust, institutional trust, and organizational trust. Yet, because the framework of proof rules for evaluation remains in the process of refinement, evaluative conclusions may not always rest on

26 For example, the *Basic Code of Professional Ethics for Judges of the People's Republic of China* provides that judicial ethics for judges include loyalty to the judicial cause, the assurance of judicial justice, the safeguarding of judicial integrity, adherence to justice for the people, and the maintenance of the judicial image. The *Basic Code of Professional Ethics for Prosecutors of the People's Republic of China (Trial Implementation)* provides that prosecutorial ethics include loyalty, justice, integrity, and civility.

27 For example, normative documents such as the *Opinions on Strengthening Governance of Science and Technology Ethics* issued by the General Office of the CPC Central Committee and the General Office of the State Council, the *Interim Measures for the Administration of Generative Artificial Intelligence Services* issued by seven departments including the Cyberspace Administration of China, and the *Guidelines on the Code of Ethics for Digital Technologies* issued by the China Association for Science and Technology.

a sufficiently clear justificatory basis when challenged. This is reflected in three respects. In this article, proof rules do not refer to evidentiary rules for adjudicating individual cases, but to rules through which evaluative conclusions are justified, explained, challenged, and reviewed.

Firstly, the facts to be proved tend to be obscured by black-box processes. What evaluation must prove are propositions concerning the state of judicial power in operation, the performance of judicial personnel, and the appropriateness of case handling. In practice, these propositions are often compressed into indicator fluctuations or model labels. Moreover, judicial business data contain a high proportion of unstructured materials and remain insufficiently structured,²⁸ while steps such as data cleansing, missing-data treatment, outlier exclusion, and the conversion of indicator rules are difficult to present in full; the grounds for model inference are likewise not easily externalized. As a result, the facts to be proved are narrowed from substantive propositions to explanations of outputs.

Secondly, the burden of justification and explanation becomes blurred. Data generation, the maintenance of indicator rules, and the use of evaluative conclusions belong to different organizational actors. Duties of disclosure, explanation, and verification are not clearly allocated. When disputes arise, it becomes difficult to determine who bears responsibility for explaining which stage, and the organization correspondingly struggles to form unified review standards and stable responsibility attribution.

Finally, objection mechanisms tend to be obscured. Evaluative conclusions are jointly produced by data-extraction processes and indicator rules. Thresholds, weights, generation logic, or model parameters are often technically encapsulated, while the front end displays only scores, rankings, or labels. Objections are therefore easily displaced into disputes over individual indicators or data points, making it difficult for challenges to reach substantive review of indicator rules and the conditions of model inference.

At a more fundamental level, the difficulty confronting judicial justice evaluation under digital and intelligent empowerment lies in the excessive compression of evaluative propositions into result-oriented expressions. Many evaluative conclusions ultimately appear only in the form of scores, rankings,

28 See Zheng Zhihang & Cao Yonghai, *The Difficulties and Responses in Quantitative Evaluation of Judicial Quality in the Digital Era*, 2 Gansu Social Sciences 128-138 (2024).

labels, or warning signals. While this seems to improve the efficiency of information transmission, it simultaneously compresses the question of which facts are to be proved, which rule conditions have been applied, and which inferential steps have been taken in generating the conclusion. Once the facts to be proved are compressed, participants in the evaluation process can no longer readily identify the true focus of dispute; once the focus of dispute becomes blurred, the burden of proof can no longer be clearly allocated among data providers, rule setters, model users, and evaluation subjects; and once the allocation of responsibility becomes inaccurate, objection handling is likewise reduced to generalized disputes over whether the result appears reasonable, rather than entering specific and reviewable levels of proof. In this way, the evaluative process becomes technologically encapsulated, proof problems recede into the background, and the credibility of evaluation ultimately loses its testable procedural foundation.

Accordingly, under conditions of digital and intelligent empowerment, the task facing the theory of proof rules is no longer merely to answer whether an evaluative conclusion has a basis. It must further determine who is to present that basis, who is to explain the governing rules, who is responsible for the inferential process, and how objections ought to be raised and examined. In other words, once judicial justice evaluation enters a data-driven setting, questions of proof are no longer confined to the determination of facts in individual cases. They extend to the authenticity of the sources of evaluative data, the accuracy of the application of statistical criteria, the adequacy of explanations for model outputs, and the reviewability of the process through which conclusions are formed. Only by bringing these matters back within the field of proof can judicial justice evaluation overcome the dilemma in which results are visible but the process remains mute, and provide a clearer analytical framework for subsequent objection review, responsibility tracing, and institutional revision.

V. THE THEORETICAL ADAPTATION OF JUDICIAL JUSTICE EVALUATION UNDER DIGITAL AND INTELLIGENT EMPOWERMENT

The foregoing three sets of theoretical challenges indicate that digital and intelligent empowerment enhances judicial justice evaluation's capacity for data extraction and computation, yet it does not automatically

generate a corresponding normative order. For data-driven evaluation to acquire legitimacy in the judicial context, the core task is to embed technological capabilities within a normative framework constituted by the rules governing judicial activities, professional ethics, and proof rules. Accordingly, theoretical adaptation may proceed on three levels: an evaluation architecture guided by the objective rules of judicial activities, the incorporation of digital and intelligent values into professional ethics, and the application of proof rules under digital and intelligent conditions.

A. Establishing an Evaluation Architecture Guided by the Objective Rules of Judicial Activities

The objective rules of judicial activities combine generality and particularity in judicial operation and should therefore inform the *ex ante* stages of indicator design and model inference, serving as the normative basis of a digitally and intelligently enabled evaluative architecture.²⁹ On this basis, establishing such an architecture requires translating the rules governing judicial activities from abstract interpretive requirements into concrete constraints that run through algorithm design, model inference, and indicator configuration.

Firstly, algorithmic design should be adjusted around the rules governing judicial activities. The structured representation of judicial phenomena within digital and intelligent systems should not stop at field mapping and statistical aggregation. It should incorporate, in a more tractable form, elements of the rules governing judicial activities such as procedural conditions, responsibility allocation, and requirements of reason-giving into data structures and inferential chains. While strengthening structured³⁰ and rule-based³¹ expression, the system should provide large judicial models with knowledge support, such as knowledge graphs, so that those elements can be incorporated into traceable inferential chains and the system can preserve retraceable rule-based starting points, thereby reducing

29 See Xie Shu, *Theoretical Inheritance and Knowledge Extension of the Structure of Criminal Procedure: From the Perspective of Cognitive Science*, 2 *Political Science and Law* 128-141 (2022).

30 See Zhang Baosheng & Yang Jing, *An Evidentiary Reasoning Model for Artificial Intelligence Legal Systems*, 5 *Evidence Science* 517-533 (2021).

31 See Tian Lizhi, *On the Stylistic Features of Chinese Judicial Opinions During the Modern Transitional Period*, 6 *Literature, History, and Philosophy* 115-125 (2012).

argumentative distortions caused by contextual compression.

Secondly, the functional boundaries of large judicial models within evaluative systems should be clearly delineated. Such models should undertake auxiliary tasks, such as generating leads, issuing risk alerts, and conducting similar case comparisons, rather than producing final and conclusive judgments, such as determinations of case-handling quality and effectiveness. Model outputs should be able to point back to specific node-level conduct and documentary materials and serve as entry points for verification and re-examination; they should not replace normative reasons with inferential labels.

Finally, the indicator system should be recoded under the constraints of the rules governing judicial activities so as to reduce conflicts among data categories, redundant repetition, and inconsistencies among counting rules.³² Indicators should also be positioned and used separately across overall evaluation, management evaluation, and specialized evaluation so as to avoid interpretive misalignment and functional overexpansion resulting from cross-scenario migration of indicators.

Moreover, overall evaluation, management evaluation, and specialized evaluation cannot be governed by a single indicator logic. This is not merely because they differ in object, but more fundamentally because they carry different evaluative functions. Overall evaluation is concerned with capturing the general state of judicial operation, and its indicators should therefore remain relatively general and stable. Management evaluation is concerned with identifying weak points in the process of duty performance, and its indicator arrangement accordingly places greater emphasis on node control, procedural verification, and the correspondence of responsibility. Specialized evaluation, by contrast, is typically directed toward particular categories of cases, particular institutional stages, or particular risk-related matters, and its judgments depend more heavily on expert knowledge, contextual understanding, and targeted review. If high-frequency and rigid indicators originally designed for managerial control are directly transplanted into overall evaluation, overall evaluation is liable to degenerate into no more than an aggregation of reports. Likewise, if matters belonging to

32 See Yu Xingzhong, *An Introduction to the Application of Predictive Coding in Justice*, 1 *Economic and Trade Law Review* 97-103 (2018).

specialized evaluation and properly dependent on professional judgment are forcibly incorporated into a uniform quantitative metric, differences among individual cases and the space for professional explanation will be unduly compressed. Accordingly, an evaluation architecture guided by the rules governing judicial activities should establish a proper correspondence among evaluative levels, evaluative functions, and the intensity of applicable rules, so that different types of evaluation remain within their proper bounds while preserving their respective emphases.

B. Supplementing Judicial Professional Ethics with Digital and Intelligent Values

Judicial professional ethics constrain official conduct through value requirements and responsibility, and provide a basis of legitimacy for adjudicative and managerial activities.³³ Under digital and intelligent empowerment, judicial justice evaluation relies more heavily on business-system data and rule-based computation. Consequently, the regulatory focus of professional ethics can no longer be confined to an adjudicative attitude and personal integrity; it must also cover the use of data and rules. Otherwise, the demand for justice may be displaced by efficiency orientation within technology-mediated evaluation.

Digital justice may be understood as a necessary development of judicial professional ethics under digital and intelligent empowerment and may be distilled into four core duties.

Firstly, a duty of data integrity should be clarified. Judicial personnel and other evaluation actors bear responsibilities to ensure that system records are truthful, complete, and timely, and must not alter the meaning of records through data-distorting practices such as selective entry, *ex post* supplementation, or strategic splitting.³⁴ Where supplementation or correction is necessary, the basis and reasons should be recorded concurrently so that records correspond to actual official conduct.

Secondly, a duty of norm-conforming use should be strengthened.

33 See Charles G. Geyh, *The Architecture of Judicial Ethics*, 8 University of Pennsylvania Law Review 2351-2399 (2021).

34 See Yao Changqing & Tian Ruiqiang, *A Study of Data Integrity in Open Science*, 1 Science-Technology & Publication 130-135 (2019).

Indicator rules and their thresholds and grading standards directly affect responsibility judgments and resource allocation. Their formulation and application should therefore remain cautious. The rules governing judicial activities should not be ignored for the sake of convenient ranking, nor should a single efficiency variable be allowed to override elements of quality and due process. Rule adjustments should preserve continuity and intelligibility so as to sustain substantive evaluation of judicial work.³⁵

Thirdly, a duty of algorithmic prudence should be established. Model prompts, automatic markings, and ranking labels should be treated with caution, and inferential outputs should not be equated with facts or conclusions.³⁶ Evaluations involving responsibility judgments and case quality should not rest solely on system labels; they should return to case context and normative reasons in order to complete the judgment and provide an account.³⁷

Finally, a duty to safeguard rights should be strictly enforced. When digital and intelligent evaluation is applied in judicial management, due regard should be given to the actual experiences of the parties. Scores and labels must not be used to curtail their rights or narrow the space for remedies, nor should evaluation results be invoked to dilute the requirements of procedural safeguards.³⁸ Where conclusions may produce adverse effects, room should be preserved for statements and explanations, so as to prevent technological conclusions from imposing disproportionate constraints on individual cases.³⁹

These duties are necessary precisely because deviations in digitally enabled evaluation often do not arise in the form of manifest illegality, but occur in a concealed manner under the appearance of smooth procedures, normal system operation, and clear results. The duty of data integrity is intended to prevent distortions in foundational facts caused by data-entry strategies, supplementary recording, and the selection of fields. The duty of norm-conforming use

35 See Liu Fuyuan, *A Study of Indicator Systems in Civil Servant Evaluation Norms: Toward the Substantive Evaluation of Civil Servant Conduct*, 5 *Theory and Reform* 110-118 (2015).

36 See Aleš Završnik, *Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings*, 5 *European Journal of Criminology* 623-642 (2021).

37 See Yuan Zeng, *A Reexamination of the Limited Legal Personality of Artificial Intelligence*, 5 *Oriental Law* 50-57 (2017).

38 See Justin Nix, Scott E. Wolfe & Jeff Rojak et al., *Trust in the Police: The Influence of Procedural Justice and Perceived Collective Efficacy*, 4 *Crime & Delinquency* 610-640 (2015).

39 See Yin Bo & Tang Shuchen, *Conflict and Coordination Between the 'Case-to-Case Ratio' Assessment and Prosecutors' Duty of Objectivity*, 2 *Rule of Law Forum* 225-238 (2021).

is intended to prevent evaluators from covertly rewriting the premises of applicable rules through technological tools, so that different standards are effectively applied under the same evaluative label. The duty of algorithmic prudence is intended to prevent system recommendations, model warnings, and automated comparison results from being mistaken for final judgments, thereby weakening the duties of explanation and review that judicial personnel themselves ought to bear. The duty to safeguard rights, in turn, is intended to prevent scores, rankings, and labels from compressing the space for explanation, objection, and remedy available to evaluated subjects, so that they are not reduced to passively accepting results without being able to enter the process by which conclusions are formed and raise effective objections. In this sense, supplementing judicial professional ethics with digital and intelligent values is, in essence, an effort to preserve the human judgment, sense of responsibility, and procedural boundaries that judicial justice evaluation ought to maintain under conditions of increasingly powerful technological operation.

C. Improving the Application of Proof Rules in Judicial Justice Evaluation Under Digital and Intelligent Conditions

The significance of proof rules lies in transforming evaluative conclusions from mere result display into a verifiable argumentative process. Although digital and intelligent empowerment substantially expands both the sources and structural forms of evaluative data, it does not necessarily increase the credibility of evaluative conclusions. The data on which judicial justice evaluation relies are markedly heterogeneous: same-source logs coexist with externally aggregated data; structured fields coexist with unstructured texts; and complete records coexist with incomplete records. Accordingly, the application of proof rules under digital and intelligent conditions should be improved through three institutional arrangements.

Firstly, rules for the admissibility of evaluative data should be established. Data admitted into evaluation should meet minimum standards of use, including verifiable sources, traceable formation processes, clear counting rules, and explainable missingness and error correction. For data that do not satisfy these conditions, their scope of use and evaluative weight should be specified, so as to avoid definitive judgments in circumstances of

data silos, inconsistent fields, or insufficient structuring.⁴⁰

Secondly, rules for allocating the burden of proof should be clarified. The chain of digitally enabled evaluation involves multiple institutional interfaces, and the burden of proof should be correspondingly allocated to enable responsibility tracing. On the case-handling side, case-handling departments and their internal lines should bear a duty of recording to ensure that data are truthful and complete and that logging standards are met. On the platform side, system builders and operators, as well as data governance and rule configuration units, should bear a duty concerning rules so that data settings, computational conditions, and version changes are documented and retraceable. On the management side, departments that use evaluation results, such as adjudication management and procuratorial management departments, should bear a duty of use to ensure consistency among evaluation purposes, the scope of effects, and the strength of conclusions. Only where boundaries of responsibility are clear can evaluation disputes be directed to identifiable objects of inquiry.⁴¹

Finally, rules for accepting and reviewing objections should be improved. An objection review procedure should be used to correct evaluative conclusions so as to better realize the governance function of judicial justice evaluation. To that end, objection handling should take a unified and clear normative form and be embedded as process nodes and document templates that can be activated within judicial business systems. Objections to judicial justice evaluation may be classified, according to their direction, into factual objections, rule objections, and inferential objections. Factual objections concern consistency between original records, case-file contents, and evaluative conclusions, and should be examined by returning to original data records and case materials. Rule objections concern the accuracy of counting rules, statistical scope, and conditions of application, and should be reviewed in light of the evaluative rules and their prerequisites. Inferential objections concern the appropriateness of model parameters, threshold settings, and grading bases, and should be examined by returning to model settings and inferential chains. With respect to

40 See Zhu Xinli & Zhou Xuyang, *Balancing the Use and Protection of Personal Data in the Era of Big Data: The Proposal of a 'Resource Access Model'*, 1 *Journal of Zhejiang University (Humanities and Social Sciences)* 18-34 (2018).

41 See Dong Binzizi, Wang Chengwei & Hua Leqin, *Replication Experiments in Public Administration: Evaluation Indicators and Research Agenda*, 6 *Public Administration and Policy Review* 149-168 (2024).

rule objections and inferential objections, technical encapsulation should not be allowed to leave only scores, rankings, or labels at the front end, thereby reducing evaluation disputes to contests over individual indicators.⁴²

On this basis, objection review in digitally enabled evaluation should also generate outcomes corresponding to the three categories of objections. In the case of factual objections, it should be clarified whether the relevant data must be corrected, supplemented, or excluded, and whether such changes have any substantive effect on the existing evaluative conclusion. In the case of rule objections, it should be determined whether there is any deviation in the evaluative criteria or conditions of application, and, where necessary, the evaluative conclusion should be recalculated or restated. In the case of inferential objections, the review should focus on whether the process of reaching the conclusion involves oversimplification, inappropriate analogy, or the substitution of model output for independent judgment, and should decide, in light of the results of the review, whether the original evaluative conclusion should be maintained, adjusted, or revoked. In this way, conclusions of judicial justice evaluation cease to be merely one-way outputs at the endpoint of management and instead become subject to a procedure in which they can be explained, challenged, and revised, thereby enhancing the acceptability and credibility of judicial justice evaluation under conditions of digital and intelligent empowerment.

VI. CONCLUSION

Judicial justice evaluation bears on the public's sense of fairness and justice in each case and the sustained formation of judicial credibility, and it also provides continuing judicial support for the modernization of the national governance system and governance capacity. Taking digital and intelligent empowerment as its starting point, this article has examined the problem of theoretical adaptation arising as judicial justice evaluation moves from traditional *ex post* statistical summaries toward data-driven governance. The key to digitally and intelligently empowering judicial justice evaluation does not lie in further expanding the indicator system, but in forming a

⁴² See Shen Shouwen, *Judicial Democracy or Judicial Independence: An Interpretation of the Nature of 'Judicial Dissent'*, 3 *Studies in Law and Business* 100-107 (2014).

theoretical framework capable of constraining technological use while continuously producing acceptable evaluative conclusions.

On this approach, the objective rules of judicial activities should operate as *ex ante* constraints on indicator design and model inference so as to prevent evaluative rules from departing from judicial procedures and the logic of adjudication. Digital justice should function as an extension of judicial professional ethics under data-based conditions by incorporating data integrity, norm-conforming rule use, algorithmic prudence, and rights protection into the operation of evaluation. Proof rules, in turn, should organize evaluative conclusions into a verifiable argumentative process through institutional arrangements concerning the admissibility of evaluative data, the allocation of responsibility, and the review of objections, thereby ensuring that evaluative conclusions can be questioned, re-examined, and corrected.

This article remains primarily at the level of analyzing paradigms and mechanisms of judicial justice evaluation. It does not yet provide a sufficiently typological examination of differences among business systems across regions and functional lines, nor of the more concrete linkage between indicator rules and model inference. Future research may move further into specific operational scenarios to test the contextual fit and operational robustness of judicial justice evaluation systems, to identify concrete pathways for embedding digital justice into judicial professional ethics and evaluative rules, and to explore more practicable mechanisms for justifying conclusions and reviewing objections. Such research can enhance judicial governance capacity while maintaining the basic requirements and boundaries of responsibility associated with judicial justice. It can also further support the regulated exercise of judicial power.

(Edited by Li Jiesi)