

A Brief Discussion on the Application of Evidence Law in the Big Data Era

Quan Li

School of Law, Southwest Petroleum University, Chengdu, Sichuan, 610500

ABSTRACT

With the advent of the era of big data, the field of evidence law is undergoing unprecedented changes. The purpose of this paper is to explore the application of big data technology in evidence law and its impact on the current legal system. First, this paper outlines the basic concepts of big data and its potential in the legal field, including improving the efficiency of evidence collection and enhancing the depth and breadth of data analysis. Then, this paper analyzes the application status of big data evidence in criminal and civil litigation, including its role in case investigation, evidence presentation and court debate. This paper further discusses the legal challenges posed by big data evidence, such as privacy protection of evidence, verification of authenticity and integrity of data, and algorithmic transparency issues. The paper proposes a range of possible solutions to these challenges, including the development of new rules of evidence, the strengthening of data protection regulations, and the improvement of judicial personnel's understanding and application of big data technologies. Finally, this paper emphasizes that in the era of big data, evidence law must adapt to technological development to ensure the legitimacy, relevance and sufficiency of evidence, so as to maintain judicial justice and improve the accuracy of legal decisions. Through a critical analysis of the existing legal framework and a prediction of future trends, this paper provides a comprehensive perspective on the application of evidence law in the era of big data.

KEYWORDS

Big data; The law of evidence; Judicial application; Data protection; Algorithm transparency

1 Definition and Characteristics of Big Data

Big data generally refers to a collection of data that exceeds the capabilities of traditional data processing, characterized by volume, velocity, variety, veracity, and value. The development of big data technology has evolved from early data warehouses to modern real-time analysis and cloud computing. With the continuous development of digital technology and the ongoing advancement of judicial reform, the determination of judicial facts in the digital age is showing six major development trends: digitalization of evidence, remote evidence collection, blockchain-based evidence preservation, virtual evidence presentation, asynchronous evidence examination, and intelligent evidence authentication. Traditional evidence law is also facing comprehensive and profound challenges due to the judicial application of digital technology: firstly, the constantly emerging types of evidence with the development of digital technology have gradually broken through the legalism of evidence categories; secondly, online litigation and asynchronous evidence examination models pose challenges to the basic principles of evidence law, such as the principle of direct oral testimony, the principle of concentrated trial, and the best evidence rule; thirdly, the complexity of digital technology makes the examination and judgment of the evidentiary attributes of relevant evidence more difficult; fourthly, the big data experience in evidence reasoning is extremely dangerous due to the integrity, fallibility of the data itself, and the algorithmic black box problem; lastly, the judicial application of digital technology has led to the weakening of evidentiary rights protection, etc. To cope with these challenges, evidence law in the digital age should move from a closed to an open evidence law, from an evidence law of information regulation to one of risk prevention and control, and from an evidence law of power regulation to one of rights protection.^[1]

2 Theoretical Basis of Evidence Law in the Big Data Era

Conviction is an extremely complex link in the criminal justice process, and the big data era poses new challenges and requirements for the theoretical basis of evidence law. In traditional evidence law theory, the collection, analysis, and presentation of evidence mainly rely on manual operations, while in the big data era, the acquisition and processing of evidence increasingly depend on technical means, especially data analysis and algorithmic models. Big data evidence, due to its scale and complexity, is increasingly highlighting its probative value in judicial practice. It can reveal patterns and trends by analyzing a large amount of data, thereby providing new perspectives and evidence for cases. At the same time, big data evidence may be regarded as a new type of evidence, standing alongside traditional physical evidence, documentary evidence, and witness testimony. It includes conclusions and reports derived from data analysis.

In the digital age, the review of the legality and authenticity of big data evidence has become more complex. It is necessary to ensure that the collection and processing of data comply with legal provisions, and that the results of data analysis are reliable and accurate. However, the relevance between big data evidence and the facts of a case may not be

as direct and obvious as that of traditional evidence. Therefore, new rules of relevance need to be established to assess the relevance and probative force of big data evidence. Generally speaking, there are four main forms of the application of big data evidence: first, using the carrier of big data as evidence; second, using the equivalent copy of big data as evidence; third, using a part of the data in big data as evidence; and fourth, using the conclusions of big data analysis as evidence. The first three forms can be covered by traditional categories of evidence, while the fourth cannot. In practice, it is mostly transformed into the form of expert opinions. Big data evidence also requires courts and judicial personnel to update their methods of reviewing and judging evidence. This may include understanding and evaluating data analysis methods, as well as demanding algorithmic transparency and interpretability. In the era of big data, evidence analysis increasingly relies on technical assistance, such as data mining, machine learning, and artificial intelligence. The application of these technologies needs to be properly regulated and guided in the law of evidence. In terms of the exclusion of illegal evidence, big data evidence may require specific exclusion rules to deal with issues such as inaccurate data, defective analysis methods, or illegal data collection. In addition, when dealing with big data evidence, it is necessary to consider the protection of personal privacy rights and data protection rights, and ensure that the collection and use of evidence do not infringe upon these fundamental rights. Scientific evidence, represented by big data evidence, brings opportunities for the transformation of criminal trials, while also having a profound impact on the traditional evidence system and rules.^[2]

With the rise of big data evidence, judicial personnel need to receive relevant training and education to enhance their understanding and application ability of data analysis technology. With the development of digital technologies such as the next-generation Internet, artificial intelligence, blockchain, and extended reality, the first-generation online courts are moving towards the second-generation online courts. Looking at the current prospects of digital technology application and the international frontier development trends, the second-generation online courts have technical characteristics such as remote presence, augmented reality, virtual reality trials, and legal artificial intelligence. In terms of international cooperation and standard setting, the transnational nature of big data evidence requires the international community to cooperate in the theory and practice of evidence law and jointly formulate international standards and guiding principles.

3 Judicial Practice of Big Data Evidence

The application of big data evidence in judicial practice is an important issue in the field of law in the digital age. Big data evidence usually comes from a large amount of electronic data, such as communication records, transaction data, and network behavior logs. These data need to be collected, stored, and analyzed through specific technical means in order to be used in judicial procedures. China's evidence legislation has a certain degree of closure in the provisions of statutory evidence categories, which is mainly reflected in Article 50 of the Criminal Procedure Law. Paragraph 1 of this article defines the concept of evidence, and paragraph 2 immediately stipulates eight types of evidence such as physical evidence and documentary evidence. However, it does not end with the word "etc.", presenting a closed statutory evidence category. Articles 63 of the Civil Procedure Law and 33 of the Administrative Procedure Law also make similar provisions. This has led to a theory in China's evidence legislation and practice that "evidence that does not conform to the statutory evidence category (form) cannot be used as the basis for conviction", which we can call "statutory evidence category-ism" or "statutory evidence form-ism".

In 2012, when the Criminal Procedure Law was amended, electronic data was added. In 2021, the new Interpretation of the Supreme People's Court on the Application of the Criminal Procedure Law of the People's Republic of China added new types of evidence such as expert reports and accident investigation reports. Second, as mentioned above, big data evidence and artificial intelligence evidence, which are not statutory evidence categories, have been widely applied in judicial practice. Big data evidence is a completely new type of evidence. Its form is different from any traditional type of evidence. Not only can it not find the basis for its validity in legal provisions, but it also cannot find a corresponding way in theoretical paradigms such as scientific evidence and probabilistic evidence. The "statutory evidence category-ism" that "whether it belongs to the statutory evidence category is one of the judgment factors of evidence qualification" is a remnant of the statutory evidence system that has long been abandoned by the civil law system. It has obviously begun to be gradually broken through in judicial practice. There is controversy over the legal positioning of big data evidence in judicial practice. Some views hold that it should be included in existing types of evidence, such as expert opinions; while other views advocate that it should be regarded as a new type of evidence. The application of big data evidence in the judiciary needs to pass the tests of relevance and scientific reliability. This involves the timeliness, integrity, and accuracy of the data, as well as the scientific and rational nature of the analysis methods. The evaluation of its probative force needs to consider the selection of information connection points, the possibility of algorithmic discrimination, the rationality of the logical structure, and the robustness of the algorithmic conclusions. In judicial practice, the analysis of big data evidence often relies on machine learning and artificial intelligence technologies, which requires judicial personnel to have corresponding technical knowledge or the assistance of experts.

The presentation of big data evidence in court may require specific technical means, such as data visualization, to help the court understand complex data relationships. At the same time, the defense has the right to cross-examine big data evidence, including questioning the source of the data, the analysis methods, and the conclusions. In the process of collecting and analyzing big data evidence, it is necessary to fully consider the issues of personal privacy rights and data protection to ensure that the basic rights of citizens are not infringed.

With the widespread application of big data evidence, the existing legal framework may need to be reformed to adapt to the new form of evidence. This may include amending the rules of evidence, enhancing the training of judicial personnel, and updating the court's trial methods.

In actual cases, big data evidence has been used as investigative clues, to refute the defendant's defense, to strengthen testimonial evidence, and to corroborate circumstantial evidence. In summary, the structure of rights and responsibilities of the main bodies should be optimized in terms of strengthening the obligations of the prosecution, safeguarding the rights of the defense, and improving the adjudication rules; the trial cross-examination rules should be refined in aspects such as updating the expert assistant system, detailing the rules for excluding illegal evidence, and clarifying the rules of cross-examination; and the drawbacks of the algorithmic black box should be overcome from aspects such as promoting limited disclosure of algorithms, enhancing the understandability of algorithms, and adding data access rights, so as to correct the cross-examination system of big data evidence.^[3]

4 Development Trends and Prospects of Evidence Law in the Big Data Era

Faced with the complex challenges of evidence law in the digital age, evidence law needs to achieve a transformation from being closed to being open. Firstly, the categories of evidence should be open. For the classification of big data evidence, a phased strategy can be adopted to gradually realize its standardized application in court. Specifically: 1. In the short term, big data evidence can be regarded as expert opinions. This approach helps to minimize the risks associated with the use of big data evidence and is also more easily accepted by judicial personnel. 2. In the medium term, legal amendments should be made to formally recognize big data evidence as an independent type of evidence. Given that the transparency of current big data technology and the openness of algorithms are still insufficient, classifying it as expert opinions in the long term may lead to misunderstandings in judicial practice.

Therefore, defining big data evidence as an independent category of evidence and establishing specific standards for its review and judgment will be more conducive to practical operation and judicial acceptance. 3. In the long run, the practice of using evidence categories as a threshold for admission should be gradually abandoned. Although in the early stages of judicial development, using evidence categories as an initial screening standard has certain significance for standardizing fact-finding, this approach does not fully align with the basic principles of evidence law. According to the spirit of evidence law, all relevant and lawful information should be considered potential evidence. Considering the high cost of obtaining evidence, excluding certain evidence lightly may affect the accuracy of fact-finding. Over-reliance on the legalism of evidence categories may lead to the formalization of the fact-finding process and weaken the legitimacy of the procedure.

Through this gradually open strategy, evidence law will be better able to adapt to the needs of the digital age while ensuring judicial fairness and efficiency. As a new form of evidence, big data evidence is gradually being incorporated into the legal framework. Future evidence law may need to clarify the legal status of big data evidence and formulate corresponding rules. With the development of big data technology, evidence law needs to innovate theoretically to adapt to the changes brought by big data. This includes a re-examination of the basic categories such as the relevance, probative force, and admissibility of evidence. The big data era demands a reconstruction of existing evidence rules to accommodate the characteristics of big data. This may involve rules regarding the collection, preservation, analysis, and presentation of evidence. Legal professionals also need to cooperate with technical experts to ensure the scientific and legal nature of big data evidence. This requires a deep integration of law and technology, as well as relevant technical training for judicial personnel.

The review and judgment of big data evidence require new methods and standards to ensure its authenticity, legality, and relevance. This may include the transparency of algorithms, the quality of data, and the reliability of analysis methods. With the use of big data evidence, judicial transparency needs to be enhanced. This includes the disclosure of the source, processing, and analysis results of big data evidence. In the process of using big data evidence, the rights of the parties, including the right to know, the right to cross-examine, and the right to defend, need to be fully protected. Big data evidence may bring new risks, such as wrongful identification and privacy infringement. Therefore, the development of evidence law needs to focus on risk prevention and control, and formulate corresponding preventive and response measures. The development trend of evidence law requires continuous academic research and practical exploration to constantly adapt to the changes in technological progress and judicial needs.

5 Conclusion

The big data era will lead the next wave of technological innovation. At present, traditional electronic data notarization mainly relies on national notarization, which is not only costly but also has a time delay from the discovery of infringement to the completion of notarization, increasing the risk of evidence loss. For example, in defamation cases handled by Internet courts, the evidence is often web pages posted by infringers on websites. According to traditional notarization procedures, victims need to invest a lot of time and money during the working hours of notary institutions to convert web page evidence into paper notarization documents.

However, with the help of Internet notarization technology, victims can upload infringing web pages to the notarization platform at any time and at a low cost, achieving instant fixation of evidence. Blockchain notarization has significant advantages over traditional notarization: it is less costly, more convenient to operate, and is not limited by time and place.

Electronic data has strengths that traditional notarization cannot compare with, and it is expected that China's civil litigation evidence system will more actively integrate blockchain technology. With the evolution of electronic data evidence preservation methods, traditional evidence theory will face challenges, and solving these challenges will help further develop big data notarization of electronic data. Through academic and legislative interpretation, various sub-items of electronic data within the scope of big data should be classified as electronic data in the statutory categories of evidence. At present, China's relevant laws have not yet clarified its classification, and legislators should clarify its status as soon as possible to lay the foundation for improving relevant evidence rules.

The authenticity of electronic data meets the standard of proof in China's civil litigation. Although it cannot be absolutely guaranteed that data will not be tampered with, the cost of tampering is extremely high, far exceeding the benefits of litigation, so the possibility of tampering only exists in theory. However, since the current technology can only ensure the integrity of data on the chain and cannot verify the authenticity of data before it is chained, when the court recognizes electronic data, it needs to combine other evidence, ensure the parties' rights to provide and cross-examine evidence, and make judgments by comprehensively applying free conviction.

Taking the pioneering case of blockchain notarization in Hangzhou Internet Court as a prime example, it is noteworthy that an extensive array of national notarization and appraisal institutions were enlisted to rigorously verify the authenticity and integrity of the data. In retrospect, certain steps in this verification process might have been deemed superfluous. This overly thorough approach can likely be attributed to the fact that it marked the inaugural instance where blockchain notarization of electronic data was formally recognized and accepted by the judiciary. Consequently, the court understandably adopted an exceedingly prudent stance, erring on the side of caution to ensure the reliability and admissibility of the novel form of evidence.

However, as the practice of blockchain notarization of electronic data continues to evolve and gain maturity, it is anticipated that the court will shift towards a more proactive and dynamic role. By embracing this cutting-edge technology with confidence, the judiciary can harness its potential to significantly enhance the quality and efficiency of civil trials. The streamlined and secure nature of blockchain notarization offers a robust alternative to traditional evidence authentication methods, reducing the time and resources required to establish the credibility of digital evidence. This not only expedites the trial process but also bolsters the overall integrity of the judicial proceedings, thereby upholding the principles of justice and fairness in the digital age.

About the Author

Quan Li, Education: Master's degree in progress, Title: Student, Research Direction: Criminal Law.

References

- [1] Zheng Fei. Where Is It Drifting: Challenges and Changes of Evidence Law in the Digital Age [J]. Local Legislation Research, 2022, 7(03): 1-18.
- [2] Hong Gang. Hidden Worries and Prevention of the Use of Big Data Evidence in Criminal Trials [J]. Forum on the Rule of Law, 2024, (02): 37-52.
- [3] Chen Zaishang, Ren Yifei. A Review of the Formalization of Cross - Examination of Big Data Evidence in Criminal Trials and Institutional Correction [J]. Journal of Henan Police College, 2024, 33(06): 52 - 61.