On the Dilemma of Examining and Admitting Online Evidence in the Digital Era and Its Regulatory Path in Evidence Law

Shuhui Wang*
School of Law, Anhui University of Finance and Economics, Bengbu 233000, China
*Corresponding email: 1446267600@qq.com

Abstract

The deep penetration of digital technologies has made the trend of "online" social life and judicial practice more prominent. As the core medium for fact-finding in the digital age, online evidence presents significant challenges to the applicability of traditional evidence law frameworks during examination and admission. This paper systematically analyzes practical dilemmas in verifying the legality, authenticity and relevance of online evidence based on the technical characteristics of online evidence. Integrating novel technological tools such as blockchain storage and electronic data forensics, it proposes regulatory pathways including legislative refinement of technical standards, judicial construction of diversified authentication mechanisms, and theoretical updates embracing a "dynamic concept of truth". The measures are intended to facilitate the coordinated development of evidence law systems and digital judicial practice.

Keywords

Online evidence, Examination and admission, Traditional evidence, Blockchain storage, Authenticity verification

Introduction

The swift digital revolution in social and economic activity has placed online evidence at the core of contemporary litigation, challenging traditional evidentiary principles like the best evidence rule and the physical truth. The article discusses the characteristics of online evidence, analyzes the main problems in the examination and admission of online evidence, and proposes optimization paths from the legislative, judicial and theoretical aspects.

The essential attributes of online evidence in the digital era and the specificity of its examination

The technical essence of online evidence and typological analysis

Online evidence is a kind of new evidence generated by digital technology. It is basically data information stored or made available by electronic means (e.g., cloud storage, blockchain nodes, smart device terminals) that can prove facts in a case. The online evidence has three features which are different from the traditional evidence (e.g., Documentary evidence, physical evidence). First, the technical dependence as its generation, storage and transmission rely entirely on the technologies such as internet, big data and artificial intelligence. Second, virtuality as its carrier is binary code (0 s and 1s), which means it has no independent physical form. Third, mutability, as data can be modified, deleted, or forged with concealed traces [1].

Based on technical generation logic, online evidence can be clearly categorized into three types. First, "native online evidence", referring to data generated entirely through digital means (e.g., social media posts, emails, smart contract transaction records). Second, "derived online evidence", which constitutes the digitized version of traditional evidence (e.g., scanned copies of paper contracts, cloud backups of offline audio recordings). Third, "hybrid online evidence", generated through the integration of online and offline processes (e.g., logistics records from e-commerce platforms involving "online ordering and offline delivery"). Different types of online evidence pose varied challenges during examination. Native evidence, detached from physical carriers, requires reliance on technical credibility for "original" status verification. Derived evidence necessitates scrutiny of both the integrity of the original carrier and the accuracy of the

digitization process. Hybrid evidence demands coordination between the spatial-temporal relevance of online and offline data [2].

The conflict between traditional evidence rules and online evidence

Traditional law of evidence is based on the concepts of "original document centrality", "testimonial evidence priority" and "physical truth". These are principles worked out from the direct perception of physical originals, from the experiential judgment of human testimony. However, in the context of online evidence, the applicability of these rules faces fundamental challenges.

(1) The decline of the "original document centrality" principle

Traditional documentary evidence favors the "original" as the best evidence and requires the parties to produce the physical original in order to determine whether it is authentic. In the world of online evidence, however, the "original" is an uncertain concept. The "original carrier" of an electronic document might be a smartphone, a cloud server, or a blockchain node - physical forms vastly different from traditional paper documents. For example, for an electronic contract sent via WeChat, the "original" can be stored in the memory of the sender device, on Tencent servers or a hash value recorded on a blockchain. Each may be considered a "original" carrier', but the legal effect of each is different, depending on the technical format [3].

(2) The limitations of the "priority of testimonial evidence" rule

Traditional evidence law is heavily reliant on witness testimony and cross-examination. With online evidence, it is the "objectivity" of the data itself that matters. For example, in an online defamation case, the plaintiff may submit a screenshot of the defamatory content on the defendant's Weibo page. The defendant can contest this evidence by saying "account hacking" or "content tampering". In the cases where authenticity is disputed, the screenshot itself is not evidence of the fact, as it only captures the surface display of a dynamic system. This has to be verified by technical evidence like tracking of IP address or login logs from the backend. This renders the traditional "priority of testimonial evidence" rule inadequate [4].

(3) The transgression of "physical truth"

Traditional evidence law aims at "objective truth". Evidence must correspond exactly to the facts of the case. However, online evidence can suffer from "technical distortion" due to technical errors (e.g., timestamp discrepancies from network latency), human tampering (e.g., Deepfake videos), or system failures (e.g., data loss from cloud servers), making it hard to get to "absolute truth" [5]. For example, blockchain storage guarantees that data cannot be changed after it is stored, but it cannot verify the truth of the raw data before it is stored. If the original data has been altered before being uploaded to the chain, the blockchain only certifies the state of the data at the time it was stored, and it cannot reconstruct the actual facts.

Core dilemmas in the examination and admission of online evidence

Legality examination: Conflicts between forensic methods and procedural norms

The legality verification of online evidence is mainly the verification of the legitimacy of the method of investigation and the observance of procedural standards. Common practices such as web crawling, on-site electronic data extraction and third-party platform data requests often lead to legitimacy disputes.

(1) Ambiguous boundaries of technical legality

Web crawling is a popular method of collecting online evidence, but its legality is disputed. Unauthorized scraping of protected data from websites is prohibited under the Cybersecurity Law and the Data Security Law and may constitute the crime of illegally obtaining computer information system data. Alternatively, the act may be lawful if the site's robots.txt protocol permits crawling (e.g., some news sites) or if the data is publicly accessible (e.g., publicly posted tweets). However, there is no consensus in judicial practice on defining "public data". For instance: Is the content of a "friends-only" WeChat Moments post "public"? If a crawler circumvents permission restrictions to access such data, is that illegal evidence gathering? Courts have reached divergent conclusions on these questions, reflecting the absence of clear statutory criteria. The legislative vacuum on these issues leads to frequent inconsistencies in judicial rulings [6].

(2) Deficiencies in procedural norms for technical investigations

In cybercrime cases, public security organs often use technical means of investigation (including on-site electronic data forensics and remote acquisition) to obtain online evidence. Such measures shall be approved strictly and shall be in accordance with statutory procedures according to the Criminal Procedure Law. But procedural norms are ill-equipped to adapt to the online environment. If a suspect's device is off during a remote inspection, can investigators still get data from a cloud backup? And if that data is on overseas servers - does international judicial assistance kick in? The Provisions on Several Issues Regarding the Collection, Extraction, and Examination of Electronic Data in Handling Criminal Cases only offer principled guidance, and are often challenged over "procedural defects" in practice.

The real problem is the opacity and lack of rigor of the current "strict approval" mechanism. This mechanism does not meet the very requirement of modern jurisprudence of external supervision over coercive investigative powers, the essence of the "Warrant Requirement". The core value of the warrant system lies in the ex ante review of coercive measures by a neutral and independent judicial authority (typically a judge). Such coercive measures include search, seizure, wiretapping and, in the online context, deep data retrieval, real-time monitoring and cloud data extraction.

This review assesses the necessity and proportionality based on "probable cause" before issuing a warrant.

First, the intrusiveness of technical investigations has increased. Remote forensics, retrieval of data from the cloud, and network monitoring can intrude far more into privacy than traditional physical searches. They involve core rights such as communication secrecy, privacy, and personal information rights. The coercive capacity of these investigations is as great or greater than that of traditional searches. The lack of independent judicial review (a warrant) and reliance only on police approval makes it hard to ensure prudence and restraint and opens the door to abuse of power.

Second, the adaptation of the "probable cause" standard is crucial. Online evidence is characterized by massive volume and ambiguous relevance. The "probable cause" standard required by warrant procedures must be interpreted in light of the digital era. When judges review warrant applications, they must assess more than the traditional clues associated with criminal elements. They must also assess the closeness of the link between the requested data scope (e.g., specific timeframes, keywords, accounts) and the alleged facts. This avoids excessive infringement of citizens' rights through "dragnet-style" data grabs - a direct response to the "massive data vs. precise proof" contradiction discussed in this paper at the collection stage.

Third, procedural legitimacy must be maintained. The issuance of a warrant by the judge is an important proof that the investigative act was legal, greatly reducing the likelihood of future disputes at trial over the validity of the procedures. Simultaneously, the warrant process compels investigative agencies to specify the scope, methods, and rationale of their requests, providing standardized guidance for evidence collection.

Therefore, it is necessary to change the current "internal strict approval" model to a substantive "judicial warrant review" mechanism to optimize the legal nature of the collection of technical investigation evidence. This evolution is not only a deepening of the Criminal Procedure Law's principle of safeguarding human rights but also a key institutional safeguard for resolving procedural ambiguities and enhancing the credibility of online evidence [7].

(3) Authority disputes in third-party platform data requests

E-commerce platforms and social media sites have collected a huge amount of user data (e.g., Taobao transaction logs, WeChat chat histories). Judicial authorities requesting such data must balance user privacy rights against the needs of criminal investigation. Under the Personal Information Protection Law, platforms that process user data usually require user consent or are covered by statutory exceptions. However, it remains unclear whether judicial data requests qualify as a "statutory exception". If a platform refuses cooperation citing "user privacy", can the court compel compliance? These questions have not been answered in the Civil Procedure Law and Criminal Procedure Law, and as a result, the prominent "platform barriers" exist in practice [8].

Authenticity verification: The chasm between technical risks and judicial cognition

Authenticity is the bedrock of online evidence, but its verification is much more complicated than that of traditional evidence. Online evidence can be secretly manipulated and forged (e.g., Deepfake videos/audio). In many cases, judicial personnel lack the technical background necessary to assess the reliability of tools such as hash algorithms or blockchain systems.

(1) The concealment of technical forgery and verification difficulties:

Deepfake technology poses one of the most significant threats to authenticity verification. There are bad actors who can use AI algorithms to swap faces and voices and generate hyper-realistic audio-visual data. For instance, in a defamation dispute, a defendant might submit a Deepfake video purportedly showing the plaintiff hurling insults. Judges relying solely on visual inspection or basic tools (e.g., resolution analysis) risk admitting such fabricated evidence as authentic [9]. Although specialized institutions have developed detection tools (e.g., Microsoft's Video Authenticator), these tools are not 100% accurate and require expert operation, limiting their widespread judicial application.

(2) Technical limitations of blockchain storage and the crisis of judicial trust

Blockchain technology, lauded for its "distributed ledger" and "immutability", is considered ideal for online evidence storage. It does have its own limitations, though. Firstly, blockchain only guarantees "immutability post-storage". It cannot verify "pre-storage authenticity". For example, if a plaintiff forges a contract locally before uploading it to a blockchain platform, the chain merely proves the file remained unchanged since upload - not its original legitimacy. Secondly, the qualifications of blockchain platforms vary widely. Technical vulnerabilities (e.g., private key leaks, node attacks) can lead to data loss or tampering, eroding judicial trust.

Empirical research on these problems has been carried out by the Research Group of Shanghai Jing'an District People's Court. In cases involving blockchain storage, approximately 30% of defendants challenged the platform's qualifications. About 20% involved disputes over discrepancies between hash generation timestamps and actual collection times. And 15% resulted in evidence exclusion due to pre-storage tampering. The numbers suggest that blockchain can enhance credibility but does not completely solve the core dilemma of authenticity verification [10].

Relevance determination: The conflict between massive data and precise proof

Relevance is the logical connection between the evidence and the facts to be proved. There is an abundance of data in the digital age, but it is not as specific as the facts of a case. A single case could involve thousands of chat logs, terabytes of cloud storage, or millions of user behavior logs. A major judicial hurdle has become filtering the data directly relevant to the disputed facts.

(1) Data overload and inefficiency of manual screening

In online infringement cases, plaintiffs can submit data, such as Weibo posts, Moments updates, and shopping records. But only a minority of it directly proves the infringement (e.g., the timestamp and dissemination path of defamatory content). Judges who rely on manual content comparison risk including irrelevant data, while over-dependence on algorithmic screening risks compromising objectivity due to "black boxes". For example, in a false advertising lawsuit against an e-commerce merchant, the plaintiff might submit product pages, user reviews, and live-stream recordings. Only the "money-back guarantee" claim on the product page is directly relevant. Isolated negative reviews may be excluded as "uncorroborated single evidence". The key data is still hard to spot effectively [11].

(2) Insufficient objectivity and explainability of algorithmic relevance

In response to data overload, some courts have turned to algorithmic models (e.g., Natural Language Processing, Machine Learning) to assist relevance determination [12]. The models automatically select relevant content by considering keyword frequency or timeline continuity. But the "black box" nature of algorithms hinders the explainability of relevance findings. Judges have difficulty grasping why an algorithm marked certain data as "relevant" or checking its accuracy. For instance, a sentiment analysis algorithm mislabeling a neutral comment as "insulting" led to an unjust verdict, sparking public concern over algorithmic fairness.

Optimization paths for evidence law regulation

Legislative level: Refining technical standards and procedural norms

Legislation is needed to clarify technical standards and procedural rules to solve the ambiguity of legality examinations. Drawing on the Provisions of the Supreme People's Court on Several Issues Regarding Case Adjudication by Internet Courts, standards for recognizing the legality of blockchain storage and electronic signatures should be detailed. Such standards should define the scope of the principle of "technology neutrality". For example, evidence obtained by illegally breaking into someone else's systems should be inadmissible. For instance, legislation could stipulate that hash values generated by blockchain storage platforms complying with national standards shall be deemed to satisfy the formal requirements for "originals". Furthermore, specific rules for emerging methods like web crawling and data scraping should be enacted via an Electronic Data Forensics Procedure Regulation. This regulation should outline pre-collection approval processes, technical certification requirements for tools, and exclusionary rules for illegal collection.

In particular, the legislation should resolve the following controversies. First, specify what is public data and state that data that is available without authentication or restrictions on access is public data and can be crawled without further authorization. Second, develop standard approval processes for technical surveys, requiring audio-visual recording of remote inspections and cloud acquisitions. These recordings should be coupled with standardized Electronic Data Extraction Records. Third, set a "necessity principle" for third-party platform requests. This principle allows for judicial demands of data only when it is necessary, e.g., when the data is exclusively held by the platform. It also protects users' right to be informed and to object.

Judicial level: Constructing diversified authentication mechanisms and technical assistance systems

To address the problem of authenticity verification and relevance determination, a two-track authentication mechanism of "technical expert assistants" and "blockchain verification" should be established.

(1) Improving the technical expert assistant system

Technical expert assistants are professionals with expertise in electronic data, blockchain, or artificial intelligence. They explain technical concepts (e.g., hash immutability, blockchain timestamps) to judges and give opinions on authenticity and relevance. While China's Civil Procedure Law allows "persons with specialized knowledge" to testify regarding appraisal conclusions, their specific role in online evidence review remains vague. Future improvements to include: First, by joint appointment of the parties in dispute, to avoid any bias. Second, to clarify the courtroom procedures, such as submission of technical reports before trial, their presentation in court with the help of visualization tools and cross-examination. Third, establishing a managed roster of qualified experts to ensure professionalism and neutrality.

(2) Optimizing blockchain technology and expanding applications

Optimization must be pursued on two fronts to overcome the limitations of blockchain. First, to improve credibility, we adopt a hybrid "consortium + private chain" architecture, and judicial organs, notary agencies and third-party platforms take part in the maintenance of nodes. Second, to develop "pre-storage authenticity verification" features. Raw data can be pre-screened before it is committed to the chain by using hash comparisons and timestamp validation. For example, the "judicial blockchain" platform of the Hangzhou Internet Court achieves "storage upon generation and full traceability". It does so by interfacing directly with e-commerce APIs to extract raw transaction data. This model, which prevents pre-storage tampering, warrants nationwide promotion.

Theoretical level: Updating the "dynamic concept of truth" and proof standards

The traditional "objective truth" paradigm in evidence law emphasizes absolute factual restoration. However, in the online world, technological constraints and data overload often render "absolute truth" out of reach. Thus a "dynamic concept of truth" must be introduced. This approach calibrates proof standards to the kind of case, the category of evidence and technical feasibility. For cases involving personal rights (e.g., online defamation) more stringent authenticity thresholds

could be imposed (e.g., by dual verification via blockchain and expert testimony). For instance, technology (e.g., blockchain) can help reduce the burden of proof in the case of property rights (e.g., online shopping disputes). For example, accepting platform transaction logs as prima facie evidence.

In addition, to handle relevance complexity, the theory of "layered relevance" can be investigated. Online evidence is divided into "core evidence" (which is directly related to the facts in issue), "supporting evidence" (which indirectly supports the core evidence), and "background evidence" (which reflects the environment in which the data is generated). Different relevance rules for each layer. For example, in an online tort case, the "core evidence" is the defendant's defamatory post. "Supporting evidence" is chat logs discussing intent to post. "Background evidence" is the defendant's account registration details. This stratification helps to avoid losing important facts in the data deluge, and increases judicial efficiency.

Conclusion

The problems of the examination and admission of online evidence in the digital age are, in essence, problems of technological rationality meeting legal rationality. These problems call for "flexible rules" to respond to "technological complexity". By refining technical standards through legislation, diversifying authentication mechanisms through judicial construction, and updating proof paradigms through theory. We can promote the collaborative development of evidence law and digital judicial practice. Only through this approach can the judicial fairness be protected and the value of online evidence be fully exploited in governing digital society.

Funding

This work was not supported by any funds.

Acknowledgements

The author would like to show sincere thanks to those techniques who have contributed to this research.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] Moussa, A. F. (2021) Electronic evidence and its authenticity in forensic evidence. *Egypt Journal of Forensic Science*, 11(20), 1-10.

[2] Das, P., Sarkar, P. (2024) The importance of digital forensics in the admissibility of digital evidence. *NUJS Journal of Regulatory Studies*, 7(2), 60-75.

[3] Luo, J. (2024) A critical review of GenAI policies in higher education assessment: a call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*, 49(5), 651-664.

[4] Horsman, G. (2022) Defining principles for preserving privacy in digital forensic examinations. *Forensic Science International: Digital Investigation*, 40, 301350.

[5] Jacobi, T., Stonecipher, D. (2022) A Solution for the third-party doctrine in a time of data sharing, contact tracing, and mass surveillance. *Notre Dame Law Review*, 97, 823.

[6] Pagallo, U., Ciani Sciolla Lagrange Pusterla, J. (2023) Anatomy of web data scraping: Ethics, standards, and the troubles of the law. *European Journal of Privacy Law & Technologies*, (2), 1-19.

[7] Serrano, M. E. (2024) The domestic and international limitations of the third-party doctrine in the digital age. *American University Business Law Review*, 13(2), 379-409.

[8] Gee, H. (2020) Last call for the third-party doctrine in the digital age after Carpenter? *Boston University Journal of Science and Technology Law*, 26, 286-300.

[9] Chesney, B., Citron, D. (2019) Deep fakes. *California Law Review*, 107(6), 1753-1820.

[10] Yusuf, A., Rahman, F. (2023) Blockchain-based evidence chains: challenges to authenticity, admissibility, and judicial trust. *Legal Studies in Digital Age*, 2(1), 53-67.

[11] Schmidt, E., Sesing-Wagenpfeil, A., Köhl, M. A. (2023) Bare statistical evidence and the legitimacy of software-based judicial decisions. *Synthese*, 201(4), 134.

[12] Cheng, H. (2022) Cost-based modeling for optimal energy management of smart buildings with renewable energy resources and electric vehicles using a scenario-based algorithm. *Advances in Engineering and Intelligence Systems*, 1(04), 14-30.