Trends, Challenges, and Responses of Legal Services in the Data Era

Authors: Shuwen Wang

Date: 2024-01-23T00:00:00+00:00

Abstract

The high activity level of data factors has created fertile ground for the development of legal data. The complexity and competitiveness of legal services have increased, and the scope of legal services has shifted from closed to open. In addition to traditional law firms, more new types of legal service providers have emerged, significantly enhancing the activity level of the legal services market. Legal data mining, due to its powerful advantages in the collection, storage, analysis, and prediction of legal data and information, is favored by some legal service providers, but the accompanying challenges and data risks also urgently need to be addressed. This reading report, based on the author' s own specialization in legal professional ethics, conducts an in-depth analysis centered on Chapter 2 of the book, "Mining Legal Data: Collecting and Analyzing the Gold of the 21st Century," which aligns more closely with the author' s research interests and professional direction, hoping to gain a more refined understanding of the chapter' s research themes. The book "Data-Driven Law" encompasses new trends in legal data such as data analysis in law firms, legal data mining, contract deconstruction, the transformation of employment big data, symbolic discourse, quantification of legal service quality, and big data revealing biases, making it highly inspiring. Chapter 2 provides three research interests: the trends and response mechanisms of legal services in the data era, challenges and risk avoidance at the level of legal professional ethics under the background of data application, and ABS, ALSP, and legal companies in the development of legal data. The analysis and research on these three themes have indeed brought much inspiration and reflection, helping to deepen the understanding of the procedural aspects of various legal service providers in the current data era regarding their establishment, operation, regulation, and evaluation, and to deepen the contemplation on the systematic construction of the new-era legal professional community and the exploration of the practice environment for new types of legal service providers.

Full Text

Reading Report on *Data-Driven Law*

This reading report, written from the perspective of legal professional ethics, offers an in-depth analysis of Chapter 2, "Mining Legal Data: Collecting and Analyzing 21st Century Gold," from Ed Walters' *Data-Driven Law*. The high activity level of data elements has created fertile ground for the development of legal data, enhancing both the complexity and competitiveness of legal services. As the scope of legal services has shifted from closed to open systems, new types of legal service providers have emerged alongside traditional law firms, significantly increasing activity in the legal services market. Legal data mining has gained favor among some providers due to its powerful advantages in collecting, storing, analyzing, and predicting legal data and information, though the associated challenges and data risks require urgent attention. This report aims to provide a more nuanced understanding of the chapter' s central themes in relation to the author' s research interests and professional orientation.

Keywords: legal data mining; new legal service providers; trends; legal professional ethics; challenges and risks

About the Book and Author

Data-Driven Law encompasses emerging trends in legal data, including law firm data analytics, legal data mining, contract deconstruction, big data transformation in employment, symbolic discourse, quantification of legal service quality, and bias revelation through big data. Chapter 2 addresses three research interests: trends and response mechanisms in legal services under the data era, challenges and risk avoidance at the level of legal professional ethics in the context of data application, and ABS, ALSP, and law companies in legal data development. Analysis of these themes offers valuable insights for deepening our understanding of the establishment, operation, supervision, and evaluation processes of various legal service providers in the data age, as well as for exploring the construction of a new-era legal professional community system and the practice environment for new legal service providers.

Chapter 2 was authored by Kenneth A. Grady, a co-founder in the legal technology field who brings a unique and timely perspective on the role of data, automation, and artificial intelligence in delivering modern, efficient services to legal consumers. As both a corporate purchaser and provider of legal services, Grady believes legal operations can achieve greater efficiency and value, particularly for corporate buyers. He serves as a visiting professor in the LegalRnD (Legal Innovation Center) program, teaching courses worldwide, conducting research, publishing, and presenting. His work focuses on improving legal service delivery and its impact on access to justice and the evolution of substantive law, with interests including artificial intelligence, attorney-client relationships, technology, legal operations management, and the future of law. Grady also works part-time at the law firm Seyfarth Shaw LLP, previously serving as CEO of

its SeyfarthLean Consulting LLC subsidiary. He is editor and primary author of the SeytLines.com blog, with numerous papers published in legal journals domestically and internationally. His industry perspectives have been featured in *The Wall Street Journal*, *Crain' s*, *The American Lawyer*, and other publications. His career includes partnership at a "Global 50" law firm, general counsel positions at three Fortune 1000 companies, partnership at a boutique law firm, and founding an independent law firm, along with senior executive roles in operations management and administrative departments at Fortune 1000 companies. For fourteen years, he has been an active member of the Association of Corporate Counsel, serving on its board, Value Challenge Steering Committee, and Advocacy Committee. He was a candidate for the 2017 College of Law Practice Management' s class and was nominated for the Fastcase 50, receiving recognition from the *Financial Times* for his innovative leadership in the in-house/outside counsel relationship.

Chapter Summary

Chapter 2, "Mining Legal Data: Collecting and Analyzing 21st Century Gold," can be structurally divided into eight parts.

Part 1: The Complexity of Legal Services describes how legal service providers once possessed geographically confined knowledge with low knowledge burdens, controllable client costs, and emphasis on long-term relationships. Today, the volume of documents, cases, materials interpreting statutes, judicial decisions, and cross-jurisdictional legal services within and across states has increased, expanding statutory scope and making legal services more complex than ever. The rising cost and burden of knowledge acquisition for providers are driving structural changes in their work.

Part 2: Competition in Legal Services explains the new changes that the data era has brought to past and present legal service entities. Previously, when data utilization was less abundant, entities with legal service needs selected providers primarily based on geographic location, access to legal information, provider pedigree, market reputation, and cost. In the current era of widespread data and technology application, selection is mainly based on cost, with brand factors becoming increasingly important. Using law firm branding as an example, a firm' s culture, track record, website, attorney qualifications, and office environment all significantly impact business development. Moreover, the previously singular competitive landscape among law firms has been disrupted by the emergence of new organizations such as legal service consulting companies and other alternative legal service providers.

Part 3: Experience-Based Advice Versus Data-Mining-Enhanced Advice distinguishes between System 1 ("experience-only advice") and System 2 ("data-mining-based advice"). System 1 represents pure "based on my experience" reasoning—fast, heuristic, but suitable only for individual providers handling small sample sizes, and inevitably subject to cognitive biases. System

2 relies heavily on data through processing and deep analysis to generate answers and add value, but is slower, cannot multitask, and struggles to interact with the world when used alone. A comprehensive approach is therefore recommended: applying System 1 when data is difficult to collect and analyze, and combining System 1 (experience) with System 2 (data) for legal data analysis and advice provision now that data collection and analysis have become easier.

Part 4: The Data Era and Legal Service Risks traces the evolution from the 19th-century Golden Age through the 20th-century Oil Age to the 21st-century Data Age, each marked by distinct characteristics. While the internet provides new data collection tools and richer analytical pathways for legal services, it simultaneously increases risks. Knowledge risk arises when providers utilize what they know or should know to deliver services, remaining relatively controllable within the provider' s sphere. However, ignorance risk exhibits uncontrollable growth, stemming from providers' failure to collect and use data that more accurately reflects reality. The expanding solution set possibilities enabled by data further increase ignorance risk.

Part 5: Integrating Data Mining with Legal Service Delivery addresses how to effectively harness the benefits of data mining and integrate them efficiently with legal service provision in the data era. It prompts providers to pay attention to the scope of data obtained through legal data mining to avoid overlooking details.

Part 6: Requirements and Components of Data Mining Programs—Data Management Plans (DMP) focuses on DMPs, emphasizing eight components of the data lifecycle: planning, collection, verification, summarization, preservation, discovery, integrity, and analysis, which together constitute the entire legal data mining process.

Part 7: Data Portability and the Value of Text Data describes how users have taken initiative over data in today' s environment, making data portable. Using data systems, legal service providers can extract data from vast materials in digital text files and parse documents by applying structure to text.

Part 8: Challenges of Emerging Data Sources and the Importance of 21st-Century Data Mining in Legal Service Provision highlights the increasing number of products with embedded sensors sending data streams to authorized recipients, raising important questions about fulfilling confidentiality obligations regarding collected client legal data. The core formula involves professional legal service providers using software to collect and analyze proprietary data, publicly available data, and client data to ultimately form services delivered to clients—constituting an important process for legal data mining.

I. Trends and Response Mechanisms in Legal Services Under the Data Era

Legal services in the data era exhibit new trends compared to traditional models, particularly regarding integration with data mining. Beyond traditional providers such as law firms, courts, and procuratorates, new legal service organizations have emerged, including legal service consulting companies, ABS, and ALSP. How these entities respond and achieve coordination and cooperation represents important content for efficiently collecting and analyzing legal data in today' s data era.

(1) Complexity and Competitiveness of Legal Services

In the past, legal services operated within narrow scopes with low activity levels. Providers' knowledge was geographically confined, creating low knowledge burdens in terms of breadth. Both clients and providers preferred long-term relationships, with clients typically selecting familiar, reliable providers within the same geographic region, resulting in controllable costs. However, as technological sensitivity has increased and data activity has grown substantially, legal activities have seen increases in documents, cases, materials interpreting statutes, and judicial precedents, with statutory scope continuously expanding. Regional barriers to legal services have gradually broken down, extending to cross-regional and even transnational scopes. These multiple factors have enriched and complicated legal services, with legal knowledge, industry practices, and standard procedures continuously expanding, imposing stricter requirements on providers.

With increased legal activity, the number and types of legal service providers have grown, intensifying competition. Beyond typical public legal service providers such as people' s courts and procuratorates, new trends have emerged among private legal service providers and organizations. Law firms now compete not only with each other but also with new legal service organizations such as legal service consulting companies, Alternative Legal Service Providers (ALSP), Alternative Business Structures (ABS), and even foreign legal service providers. The scope of legal service providers and organizations differs: the former refers to lawyers and non-licensed individuals providing legal services to clients, avoiding the pejorative term "non-lawyer" and the increasingly irrelevant task of parsing what constitutes "unauthorized practice of law." The latter includes all entities that form part of the legal service supply chain.[1]

(2) Legal Data Mining

Legal data mining, in essence, is the entire process of collecting and analyzing legal data, primarily encompassing the structure, content, and processes of materials collected and created when providers deliver services. After defining legal

[1]See Ed Walters, *Data-Driven Law*, Boca Raton: CRC Press, 2019, p. 50.

data, the preprocessing stage extracts correct, reliable, and required data from raw legal data before entering the core mining algorithm. This algorithm filters needed expression and interpretation types from numerous pattern collections, with final data conveyed to users after analysis.

Traditional online legal services mainly consist of legal database models for statutes and cases and lawyer personal marketing models (such as online consultation), neither of which leverages massive data for value discovery.[2] In the new data era, the first model can analyze and mine the application of laws and regulations in past cases, judicial logic in judgments, universal adjudication rules beyond individual case personalization, and overall win rates for certain case types. The resulting legal analysis products can be sold to law firms, making them both users of platforms and systems and providers of specific content. This analytical approach also helps strengthen supervision of public judicial authorities and promotes transparency in public legal services and efficiency in private legal services. The second model analyzes target clients' legal needs and recommends suitable lawyers, shortening demand cycles, promoting legal resource mobility, and providing clients with expected commissioning choices. Data mining technology can discover potential, easily overlooked data information in massive datasets, enabling systematic collection, extraction, and analysis amid "data explosion" to capture high-quality data.[3]

(3) Response Mechanisms of Traditional Legal Service Providers

In the era of data eruption, how traditional legal service providers adjust themselves and respond to challenges and competition from new legal service organizations warrants investigation.

1. Public Legal Service Providers

In August 2013, President of the Supreme People' s Court Zhou Qiang first proposed the concept of "big data, big pattern, big service" at the Fourth National Judicial Statistics Work Conference of the Supreme People' s Court, launching China's judicial big data initiative. The construction of China Judgments Online and China Court Trial Online places trial proceedings and judgments under network data supervision, enhancing judicial credibility and enabling timely error correction. Legal data flows have become more aggregated and enriched, with sharing mechanisms for core legal data such as judgments and trial information established and improved, and judicial informatization construction vigorously advanced.

In the construction of digital and information-based public legal services, the

[2]Hu Ling, "The Impact of the Rise of Big Data on Legal Practice and Theoretical Research," *Journal of Xinjiang Normal University (Philosophy and Social Sciences Edition)*, 2015, Issue 4, pp. 108-113.

[3]See Zhongping Xu, Xu Zhongping, Zhang Tao, Zheng Rongrong, Li Xiangna, Li Dong, "Knowledge Discovery and Analysis of ICT Customer Service Data Relying on Multi-frequency Mode Data Mining," *Journal of Physics: Conference Series*, 4(2020).

smart court initiative serves as exemplary content. The 2016 *National Informatization Development Strategy Outline* included "smart court" construction in the national informatization development strategy. The same year' s *13th Five-Year National Informatization Plan* explicitly stated: support "smart court" construction, implement electronic litigation, and build and improve the fair justice informatization project. From purely manual case file organization and information archiving to integrated informatization and manual processing in case acceptance, trial, enforcement, and supervision, and from paper-based case files and in-person trials to electronic case file recording, synchronized live broadcasting of internet trials, and data visualization of case handling processes, judicial information has become more publicly supervised, providing more stable guarantees for judicial fairness. At the machine learning level, strict requirements must first be imposed on the standardized format of input files to ensure accurate correspondence and recognition by legal analysis systems and platforms before subsequent data collection and analysis. This big data analysis method is primarily used in trial assistance systems for data extraction, tracing, monitoring, prediction, and summarization. Regarding legal data analysis tools used by public legal service providers, different tools are selected based on needs. Data extraction tools edited with Python are suitable for automatically extracting data information from websites such as China Judgments Online and can perform certain lexical analyses, offering flexible data collection methods but requiring programming skills. Excel remains a traditional tool for organization, recording, and collection, while SPSS statistical analysis tools are widely used for statistical data analysis but require standardized, structured formats for system recognition. Procuratorate data construction has also progressed alongside the deepening development of smart courts. Public legal services in the data era have effectively broken through the inherent path of judicial administrative reform, using information technology and data system platforms to open data ports and achieve mutual sharing and connectivity of legal information, integrating multiple legal services into multi-portal, continuous "one-stop" services that provide convenience for the people.[4]

2. Private Legal Service Providers

More active data elements create widening gaps among private legal service providers. Astute providers occupy more complete data resources, with more mature legal data systems, platforms, and technologies, richer cross-domain data resources, and greater capital and teams for investment. In contrast, small law firms struggle to independently build legal data systems or platforms, particularly under today' s rapidly accelerating data development trends, making survival increasingly difficult through manpower alone and ultimately leading to dissolution or merger with larger firms.

Clients constitute a crucial link in law firm business development. When data elements were less active, law firms operated within relatively closed business

[4]See Gong Yan, "System Construction and Practice of 'Internet +' Public Legal Service in the New Era," *Modern Management*, 5(2019): 634-642.

scopes focused on internal workflows. As data flows rapidly across regions, law firms now pay greater attention to clients, case sourcing, legal data, and lawyer evaluation. Using outlier data analysis methods, target clients can be efficiently identified. Outlier data analysis essentially groups similar items together. When grouping clients, multiple measurement factors such as business scope alignment and subject matter value are considered to formulate precise strategies. First, required client data is obtained and simple data cleaning is performed to filter important data for matching with law firms and specific lawyers. Second, multi-dimensional feature vectors are constructed based on client characteristics. Finally, input data is checked for structural compliance before building models to search for target clients. Conversely, on the other side of outlier target client selection, the construction of lawyer evaluation systems by clients, law firms, and regulatory agencies also occurs.

Lawyering is fundamentally a service profession providing publicly oriented legal services, making service quality evaluation crucial for both current service delivery and subsequent client selection. Therefore, constructing an industry evaluation system holds significant practical meaning. Although specialized professional rating systems for lawyers already exist, their narrow scope can no longer meet rising multi-faceted evaluation demands, necessitating the establishment of a big data-based lawyer evaluation system including authoritative lawyer identity authentication, business scope systems, client evaluation systems, and past case representation evaluation systems to achieve intuitive supervision and encourage diligent, standardized practice.

In the data era, private legal providers have transitioned from purely manual legal information collection to integrated information collection and analysis combining internet platforms, systematic legal data mining, and personal experience. This shift stems not only from the broader era but also from the genuine convenience and efficiency that data applications bring to private legal services. Human memory and concentration are inherently limited; when cases are complex and involve excessive legal elements, structured data input ensures clear organization during legal data analysis and reduces statistical and analytical errors. This is particularly evident at the retrieval level, where statutes, cases, and documents can be searched and applied more comprehensively and conveniently. At the legal document level, "artificial intelligence technology enables machine recognition of account books, invoices, and contract texts to convert paper archives into electronic documents, facilitating lawyers' drafting of legal documents and provision of legal opinions for clients through big data panoramic information recording." [5] By collecting, retrieving, and applying legal data, lawyers can also break through to major, difficult, and complex cases beyond their familiar areas, grasping macro patterns through sorting out similar past cases and combining them with specific case conditions for more efficient

[5] "Using artificial intelligence technology, through machine recognition of account books, invoices, contract texts, etc., to convert paper archives into electronic documents, and through panoramic information recording of big data, to facilitate lawyers in drafting legal documents and providing legal opinions for clients."

case familiarization and analysis to better serve clients. Big data technology also breaks through limitations in sample collection, vastly expanding scope, with hidden information helping lawyers overcome thinking blind spots and shifting their reasoning from deductive to inductive and analogical, and from reliance on causal relationships to seeking correlations, contributing significantly to case prediction.[6] Beyond basic legal services provided by law firms and lawyers, new legal services enabled by data collection, retrieval, storage, and analysis functions have entered the field. Large law firms, backed by strong capital and resources, provide solid support for developing such new legal services.

3. Legal Service Systems and Platforms

Legal service systems and platforms represent important pathways for liberating certain human resources in the data era. Common systems and platforms such as Beida Fabao, China Judgments Online, and China Court Trial Online are well-known but primarily focus on retrieval functions, rarely offering new functions like data analysis, prediction, process services, or legal resource integration and scheduling. The relationship between innovative internet legal service systems or platforms and clients resembles that between producers, sellers, and consumers, covering target client and service provider matching, process evaluation guarantees, parameter updates, and technological innovation during platform operation.

In the operation of legal service systems, mining legally structured text constitutes core content—the first and most important step in obtaining legal information. Through appropriate algorithms for individual cases, such as C4.5,[7]

[6] "The application of big data technology can also break through limitations in sample collection, greatly expanding the scope of sample collection. The implicit information can help lawyers break through thinking blind spots, promoting their thinking to shift from deductive reasoning to inductive reasoning and analogical reasoning, and from relying on causal relationships to seeking correlational relationships, making significant contributions to case prediction."

[7] C4.5 classification tree is the most popular decision tree algorithm, a supervised learning method that, given samples with attributes and predetermined categories, learns a classifier that can correctly classify new objects.

k-means,[8] Apriori,[9] EM,[10] PageRank,[11] and Adaboost,[12] the full value (or at least required value) of information in legal documents is mined, with each step contributing to ultimate thorough analysis of legal texts. Text mining technology overcomes past limitations of manual keyword searches, eliminating false positives from too many irrelevant case records to automate the process and overcome existing methodological difficulties.[13]

Access to internet legal service platforms represents a crucial link in legal data construction, determining the quality of legal service providers accessible to clients and subsequently affecting the quality and evaluation of entire legal service delivery processes. Therefore, platforms must clarify behavioral norms for law firms, lawyers, and clients, strictly review qualifications for law firm and lawyer admission, specify lawyers' practice qualifications and business expertise areas, and deny admission to lawyers with disciplinary records or violations—this constitutes pre-admission control. Regarding process evaluation and supervision, while ensuring personal information security, macro-level recording and supervision of non-confidential behaviors of each client and practicing lawyer should be conducted to establish personalized integrity files, achieving basic information symmetry among clients, lawyers, and law firms during legal service acquisition and provision to demonstrate fairness and prevent integrity risks. Additionally, by improving legal community interaction circles within legal service platforms, real-time communication between lawyers, law firms, and clients can be promoted to enhance mutual trust and increase client stickiness to internet legal service platforms.[14]

[8]The k-means algorithm, also known as k-average or k-means, is the most widely used clustering algorithm, using the mean of all data samples within each cluster subset as the cluster representative point. Its main idea is to iteratively divide datasets into different categories to optimize clustering performance criteria functions, generating compact intra-cluster and independent inter-cluster results. This algorithm is unsuitable for discrete attributes but effective for continuous data clustering.

[9]The Apriori algorithm is one of the most influential algorithms for mining frequent itemsets of Boolean association rules, involving two steps: first, iteratively retrieving all frequent itemsets in transaction databases (itemsets with support not less than user-defined thresholds); second, using frequent itemsets to construct rules satisfying user minimum confidence.

[10]The Expectation-Maximization (EM) algorithm finds maximum likelihood or maximum a posteriori estimates of parameters in probabilistic models depending on unobserved latent variables, commonly used in data clustering for machine learning and computer vision.

[11]PageRank, or webpage ranking, is Google' s proprietary algorithm for measuring the importance of specific webpages relative to other webpages in search engine indexes, reflecting relevance and importance and frequently used to evaluate webpage optimization effectiveness.

[12]Adaboost is an iterative algorithm whose core idea is training different classifiers (weak classifiers) on the same training set and combining them into a stronger final classifier (strong classifier). Adaboost changes data distribution based on classification correctness of each sample in every training set and overall accuracy of previous classifications to determine sample weights, sending modified-weight datasets to lower-level classifiers for training and finally fusing all classifiers into a final decision classifier.

[13]See M.F.M Firdhous, "Automating Legal Research through Data Mining," *International Journal of Advanced Computer Sciences and Applications*, 6(2011).

[14] "By improving the legal community interaction circles within legal service platforms, real-time communication between lawyers, law firms, and clients can be promoted to enhance mutual trust and increase client stickiness to internet legal service platforms."

Taking Faxiaobao as an example, this intelligent Q&A product provides multiple services including legal consultation, legal knowledge queries, fee calculation, document templates, and manual consultation.[15] For legal and manual consultation, cases are diverted by distinguishing case types and setting multiple channels, with initial intelligent Q&A resolving simple questions the platform can answer independently and filtering advanced questions for human transfer, efficiently allocating legal resources and reducing cumbersome manual burdens. For document templates, basic formats for various legal documents including complaints, answers, agency statements, and contracts are provided for user reference. For fee calculation, structured text mining automatically identifies and corresponds to various fees, which are then calculated according to machine-learned calculation standards within the system. "Third-party legal service platforms use data analysis and statistics to collect user information from both supply and demand sides, integrating information such as users' areas of interest, usage habits, and lawyers' professional directions to push highly relevant lawyers and clients to each other, thereby achieving an efficient matching model using data to find targets." [16]

(4) Business Boundaries Between Lawyers, Law Firms, and New Legal Service Providers

Data development has introduced competitiveness to legal services, challenging traditional law firms with new legal organizations such as legal service consulting companies, ABS Alternative Business Structures, ALSP Alternative Legal Service Providers (such as legal research services), and even foreign legal service providers. Legal service delivery will shift from a model primarily focused on internal workflows to one centered on client outcomes. Law firms will evolve from service-based hourly billing models toward product subscription models, bringing growth, high-margin sustainable recurring revenue, and new higher-value service opportunities.[17] Law firms possess vast amounts of data but must now consider how to respond to and utilize data development trends. Data-technology integration will enable new legal service providers to offer new forms of productized products, driving new recurring revenue and higher-value business consulting services.

When facing new legal service providers, must traditional law firms and lawyers optimize themselves timely? What are the overlapping scopes and boundaries in business provision among law firms, lawyers, and new legal service providers? How can lawyers' public nature and corporate profit-seeking be ensured to achieve efficient integration without conflict in business provision and coopera-

[15] Faxiaobao provides intelligent Q&A products covering legal consultation, legal knowledge queries, fee calculation, document templates, and manual consultation.

[16] Zhang Mingxin, Li Guanye, "New Developments in China' s Legal Service Industry in the Big Data Era," *Journal of Jiangsu Normal University (Philosophy and Social Sciences Edition)*, 2020, Issue 6, pp. 93-103.

[17] See Simon Drane, "How Data Will Enable the Shift Towards the Productisation of Legal Services," *Legal Information Management*, 3(2019): 138-142.

tion? These questions warrant discussion and require clarification in practice.

Lawyers practicing in law firms cannot sign labor contracts with new legal service providers such as legal service consulting companies in their capacity as lawyers, but may sign labor contracts with them in a non-lawyer capacity, as required by civil laws such as the *Labor Law of the People' s Republic of China* and *Labor Contract Law of the People' s Republic of China* that do not prohibit dual labor relationships. Using legal service consulting companies as examples, lawyers practicing in law firms may serve as legal advisors, providing business-scope services including daily legal consultation and case agency. These restrictions on lawyer practice primarily stem from Articles 10[18] and 12[19] of the *Lawyers Law of the People' s Republic of China* regarding practice venues and part-time lawyers. Lawyers may practice only in one law firm and cannot practice simultaneously in multiple venues. Part-time lawyers are exceptional, as they hold dual identities as lawyers and personnel engaged in legal education and research in higher education institutions or research organizations, meeting the conditions stipulated in Article 5[20] of the *Lawyers Law of the People' s Republic of China*—the only exception allowing dual professional identities. Overlapping business boundaries among law firms, lawyers, and new legal service organizations also involve establishing cooperation mechanisms. New legal service organizations, potentially involving more business, financial, data, and pharmaceutical fields, have broad case source demands and scopes and can introduce cases to law firms and lawyers for agency, achieving win-win cooperation. Lawyers may establish legal service consulting companies but, based on their professional public nature, cannot serve as shareholders or legal representatives (involving certain profit-seeking that conflicts with public nature), and cannot practice as lawyers therein based on the same practice venue regulations.

II. Challenges and Risk Avoidance in Legal Professional Ethics Under Data Application

In the data era, enhanced information activity and expanded distribution ranges have brought numerous challenges and risks at the level of legal professional ethics, making their response and avoidance important topics for discussion.

[18] *Lawyers Law of the People' s Republic of China*, Article 10: "A lawyer may practice in only one law firm. A lawyer changing practice institutions shall apply for a new practice certificate. Lawyers' practice is not restricted by geographical limitations."

[19] *Lawyers Law of the People' s Republic of China*, Article 12: "Personnel engaged in legal education and research in higher education institutions and research organizations, meeting the conditions stipulated in Article 5 of this Law, may apply for part-time lawyer practice with their unit' s consent and according to the procedures stipulated in Article 6."

[20] *Lawyers Law of the People' s Republic of China*, Article 5: "To apply for lawyer practice, one shall: (1) uphold the Constitution of the PRC; (2) pass the national unified judicial examination; (3) complete one-year internship at a law firm; (4) have good character. Lawyer qualification certificates obtained before the national unified judicial examination have equal effect as national unified judicial examination certificates when applying for lawyer practice."

(1) Confidentiality Obligations Regarding Client Data, Privacy Leakage Risks, and Responses

Lawyer confidentiality obligations[21] require lawyers to not disclose to third parties client data, privacy, or trade secrets obtained during practice without client consent. During legal data transmission, storage, and modification, massive legal data streams converge and require orderly sorting and storage. Regarding client metadata, which supports functions such as indicating storage locations, historical data, resource discovery, and file records with strong organization, dynamic electronic records of each client' s legal data and information are formed. However, these dynamic data records inevitably contain large amounts of legal data with implicit client data tags, making classified encrypted storage and avoidance of client privacy leakage risks an important aspect of legal professional ethics.

Data mining technology allows using massive data to discover previously unknown, potentially useful knowledge and hidden information and value, with the high speed and breadth of information technology development increasing its power and potential. However, personal sensitive information privacy is generally not respected in this process, creating ethical and legal issues in some applications.[22] Part of web mining technology aims to analyze user behavior to continuously improve website structure and content, but using powerful processing tools provided by web mining may threaten user privacy. Current legal research on privacy issues proposes a comprehensive approach based on web mining' s unique attributes that can identify behaviors potentially threatening personal privacy in each specific context to determine which technologies and uses harm client privacy.[23] But how do we define client privacy? In legal data mining, client privacy includes personal information and materials, specific content of commissioned matters, case process information for involved cases, and trade secrets for corporate clients—overall, a relatively broad scope requiring significant confidentiality obligations.

Response models for client data confidentiality and privacy leakage risks can attempt the following: First, establish collection standards and privacy statements before data collection. Miners should formulate corresponding standards before collection, clarifying which user data and information will be collected for clients to decide whether to commission services or adjust data collection scope based on case and client nature particularities. After determining collection scope,

[21] *Lawyers Law of the People's Republic of China*, Article 38: "Lawyers shall keep state secrets and commercial secrets learned in practice and shall not disclose clients' privacy. Lawyers shall keep confidential information about clients and others that they learn in practice and are unwilling to disclose, except when clients or others are preparing or committing crimes endangering national security, public security, or seriously endangering others' personal safety."

[22] Mehmet Cudi Okur, "On Ethical and Legal Aspects of Data Mining," *Journal of Yasar University*, 11(2018).

[23] Juan D. Velásquez, "Web mining and privacy concerns: Some important legal issues to be consider before applying any data and information extraction technique in web-based environments," *Expert Systems With Applications*, 13(2013).

possible fields or uses for collected data in subsequent processes must be clearly disclosed to clients, with further details subject to communication. Second, after establishing standards and procedures, legal service providers should explicitly disclose subsequent client data analysis and processing procedures to clients during the collection process, enhancing transparency and providing assurance and trust. Third, legal service providers must properly safeguard collected data, fulfill confidentiality obligations, and prevent third-party access and information leakage without violating national laws and regulations. "When data sources contain client identity, health status, property, credit, and other information, web mining software should consider assigning each client an encrypted identifier. System input should be anonymous data, with conversion mechanisms from anonymous to specific client data known only to the website. When data sources contain specific client data enabling unique identification, web mining systems should avoid using individual specific data and instead use aggregated client data, mining only aggregated data. Additionally, websites should establish a privacy control framework building a firewall between client privacy data and other data, creating mechanisms to identify and filter all user-related data in data warehouse logical models, allowing post-mining result inspection." [24] These privacy measures avoid client privacy leakage risks from the source and throughout the process, fulfilling confidentiality obligations. Recently, relevant departments have conducted window guidance for network data companies according to the *Data Security Law of the People' s Republic of China*[25] and other laws and regulations. For example, the Jufa Cases website has removed personal information of parties in judgment documents, as well as case numbers, trial courts, and judges, while other main text content remains unaffected—this also represents privacy processing for legal data information, though the protected subjects are not law firm or new legal service provider clients but rather third-party privacy in data systems or platforms.

(2) Attempts to Break Data Connectivity Barriers from a Law Firm Management Perspective

In the data era, collecting, analyzing, and utilizing legal data holds important value for legal professional community practice. Building paths for legal data interconnectivity within specific legal professional communities requires breaking necessary data connectivity barriers and opening data interaction interfaces based on actual conditions. "Current legal technology services need to establish a system, such as creating databases encompassing domestic and foreign litigation cases to promote exchange and connectivity among domestic and foreign

[24] Cao Yuping, "Research on Information Ethics Issues in Data Mining," *Information Studies: Theory & Application*, 2008, Issue 5, pp. 695-698.

[25] *Data Security Law of the People' s Republic of China* was adopted on June 10, 2021, at the 29th session of the Standing Committee of the 13th National People' s Congress and took effect on September 1, 2021. This law aims to regulate data processing activities, ensure data security, promote data development and utilization, protect individuals' and organizations' legitimate rights and interests, and safeguard national sovereignty, security, and development interests.

law firms, lawyers, and legal resources." [26] Law firms also need to build internal platforms and models for data interconnectivity, such as various contract templates formed within firms and the degree of data connectivity achievable between different groups at the same level. To balance confidentiality obligations to clients and internal information sharing, a model can be adopted where members of project teams at the same level are aware of client data for which they are responsible but maintain confidentiality regarding other teams' client data, submitting it to common superiors for encrypted analysis and processing to extract valuable, non-confidential standardized information for sharing among project groups. During sharing, only standardized, anonymized, templated data with reference value is accessed, without infringing on specific named client data and privacy. This enables learning and reference to improve law firm efficiency and achieve win-win outcomes for law firms, lawyers, and clients.

(3) Maintenance and Promotion of Legal Data Systems and Platforms

In constructing legal data systems and platforms, maintenance and promotion are crucial processes. Hidden errors discovered during use must be corrected, and new functions and performance requirements proposed by users must be expanded. Regarding data maintenance, timely updates to client data changes are needed, with proper data storage, backup, and recovery, and regular inspection of legal database operational status to ensure overall stable operation and good user experience. Similar to lawyer business promotion rules, law firms or new legal service organizations must follow certain rules when promoting legal data systems. After establishing their own legal data systems or platforms, specific law firms can use association rules and outlier data to find target clients for effective promotion, expanding audience reach and mining potential clients, which also requires system precision and client matching. Additionally, in practice, dependency on specific entities must be reduced to create inheritable, more independent operating systems to prevent vacancies and system paralysis.

III. ABS, ALSP, and Law Companies in Legal Data Development

Compared to traditional law firm development models, the emergence of new legal service providers in the data era brings new development opportunities. ABS, ALSP, and law companies are all competitive products of legal data. How these models have developed abroad and their compatibility with China' s localized development conditions are explored in this chapter.

(1) ABS

Lawyers, as a public profession, must be distinguished from general profit-oriented commercial activities. However, lawyers' businesses can indeed be

[26] Jin Feng, "Marriage with Technology May Reconstruct Legal Service Experience," *Science and Technology Daily*, June 29, 2021.

improved through active economic, resource, data, and information factors in the legal market. If market principles are fully implemented or market factors become overly active, will lawyers lose their professional characteristics in competition with non-professional legal personnel, or will they lose their original public-spiritedness in this impetuous economic society and drift toward profit-seeking? This represents a profound antinomy in the legal market. Particularly prevalent in Western developed countries are ABS (Alternative Business Structure) and legal service retail stores (such as Axess Law[27]), which commoditize legal services. While this trend indeed makes legal services more accessible to the public and expands business quantity and scope, whether such commoditization will impact lawyers' public nature or pressure law firm systems and forms to change warrants deep consideration. ABS is the abbreviation for Alternative Business Structure, adopting a business model that allows non-lawyer talent to own or invest in law firms and provide orderly legal services, permitting cooperation between lawyers and non-lawyers in legal service business ownership to create greater opportunities for all relevant parties.

The UK leads in ABS law firm establishment. As of October 2012, the UK' s Solicitors Regulation Authority (SRA) had granted four ABS law firm licenses. The SRA' s independence lies in its council members not simultaneously serving as members of the two guilds, with membership comprising both lawyers and non-lawyers in a ratio slightly favoring lawyers,[28] ensuring absolute professionalism in ABS law firm establishment personnel admission while incorporating non-lawyer professional opinions. However, ABS law firm establishment in China remains a process full of unknowns and difficulties. Due to the multiple identities of lawyer and non-lawyer personnel, whether the Ministry of Justice and judicial bureaus have sufficient grounds to regulate ABS law firms remains highly controversial. The establishment of ABS law firms in China requires theoretical and practical discussion regarding regulatory authority attribution, admission, and management system selection. Hainan Special Economic Zone has made China' s first attempt in this regard, with Article 18, Paragraph 2 of the *Hainan Special Economic Zone Lawyers Regulations*[29] relaxing admission conditions for special general partnership law firm partners, allowing certified public accountants, certified tax agents, certified cost engineers, patent agents,

[27]In 2012, Toronto lawyers Lena Koke and Mark Morris founded Axess Law, located in brand retail stores within Walmart supermarkets. Operating daily from 10 AM to 8 PM, the bright orange-themed stores provide various legal services including real estate transfers, bequests, probate, agency, and notarization. Their "instant bequest" service allows customers to obtain a sealed, effective document for $99 after waiting approximately 45 minutes to one hour, aiming to provide accessible legal services for anyone at any time with transparent pricing.

[28]Cheng Youyan, "Changes in UK' s Justice System: Non-Lawyers Owning Law Firms," *China Lawyer*, 2013, Issue 4, p. 103.

[29]*Hainan Special Economic Zone Lawyers Regulations*, Article 18, Paragraph 2: "Certified public accountants, certified tax agents, certified cost engineers, patent agents, and other professionals may become partners in special general partnership law firms, but their capital contribution shares and number proportions shall not exceed 25%, and they may not serve as law firm responsible persons."

and other professionals to become partners (with restrictions on capital contribution and number of partners not exceeding 25% each), while ensuring lawyer professional judgment independence. However, since the regulation' s promulgation, no information about specialized ABS law firm processes has emerged, indicating a long road ahead for realization in China.

(2) ALSP

Alternative Legal Service Providers (ALSP) refer to new enterprises providing legal services beyond traditional law firms. These enterprises generally do not adopt partnership systems, and management personnel need not be lawyers. Compared to traditional law firms, their corporate forms are more flexible and can compete with law firms' non-experiential businesses.[30] ALSPs are mainly divided into two categories: first, traditional enterprises providing legal services across boundaries, represented by the Big Four accounting firms' legal departments and some investment banking institutions. Due to their close cooperation with law firms or enterprises in business, their legal services overlap significantly with law firms' non-litigation businesses, and they have accumulated certain connections and legal resources through long-term practice, forming competitive relationships with law firms in non-litigation legal service provision. The second category comprises technology-based legal service companies that have seized the favorable trends of technology and data development, holding advantages in legal data collection, storage, analysis, and prediction. Compared to ABS, ALSP organizational forms are more flexible, ranging from Big Four legal departments and rapidly growing new enterprises like Legal Zoom and Elevate to small and medium-sized enterprises providing legal services to consumers and small businesses through applications and cloud services. ABS focuses on allowing non-lawyer personnel to become law firm founders. ABS law firm establishment and regulation require specific admission rules, appropriate regulatory authorities, and specific regulatory restrictions, making practical implementation more difficult than flexible ALSP forms, especially without precedent in Chinese practice (Hainan' s pilot has not made significant progress), though they have developed well in the UK, Australia, Singapore, and other countries. China is currently more suitable for ALSP and ABS structural development models, with ABS law firm development yet to be explored.

(3) Law Companies: Comprehensive Pan-Legal Service Providers

Law companies, compared to ALSPs, are more like comprehensive pan-legal service providers, encompassing accounting, IT, human resources management, and other law-related services that ALSPs lack. They can provide flexibly deployable lawyers, such as temporarily replacing employees on maternity leave in corporate legal departments or providing legal services tailored to startup devel-

[30]Li Mai, "ALSP is Booming, Where is the Legal Service Market Heading?" https://zhuanlan.zhihu.com/p/93800774, accessed November 5, 2021.

opment stages.[31] Essentially, there is no difference between law companies and law firms in providing legal services, only in the types and methods of service provision, with the former having a broader scope. Therefore, law companies do not wish to be labeled as alternative legal service providers, as "alternative" seems to imply another legitimate method (i.e., traditional law firms) while other methods are controversial and unknown. Such comprehensive pan-legal service providers actually have broad markets, particularly in law-related human resources provision, where data quickly matches suitable human resources targets to fill vacancies in legal professional positions and law-related markets.

Note: Figure translations are in progress. See original paper for figures.

[31] Zhihe Dongfang, "Law Company, the Next Era of the Legal Industry?" https://baijiahao.baidu.com/s?id=1622617017600389543&wfr=spider&for=pc, accessed November 6, 2021.