# Quantifying Legal Service Quality: Data Science Insights from the Avvo Legal Q&A Forum

**Authors:** Conghui Bi

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

## Abstract

none

## Full Text

## Preamble: Course Report

**Course Title:** Legal Services and Judicial Management Technology – Legal Professional Ethics
**Report Topic:** Quantifying Legal Service Quality: Data Science Lessons from the Avvo Legal Q&A Forum
**Authors:** Nika Kabiri, Ed Sarausad, and Rahul Dodhia
**Content Summary and Reflection**

## Chapter Summary: Quantifying Legal Service Quality

Chapter 7, titled "Quantifying Legal Service Quality: Data Science Lessons from the Avvo Legal Q&A Forum," outlines key lessons and best practices for collecting, cleaning, analyzing, and visualizing legal data. The chapter provides an overview of how to gather and store law-related data before exploring various ways such data can improve legal practice. Specifically, it examines how a legal entity might use data to evaluate lawyer quality and identify markets with high demand but scarce legal representation. Legal data sources encompass several dimensions: data acquisition, data management, data integration, data availability, and data operations.

Building upon these data sources, the second section addresses the use of data to assess lawyer quality. The context is that many individuals struggle to determine which lawyer to hire, facing a fundamental reality: people need information about attorneys yet lack systematic evaluation frameworks for such information. This leads to the chapter's central research question: how can we

evaluate lawyers to find the best ones? Two approaches emerge: a purely objective method and a hybrid objective-subjective method. The objective approach provides "parameters" about lawyers, such as years of experience, number of awards, publications, and law school rankings. Clients review these parameters much like they would when purchasing a computer or car, allowing them to see lawyer information intuitively—similar to how consumers examine product specifications on e-commerce platforms like Taobao (in this case, the Avvo forum). The hybrid approach incorporates both these objective parameters and client opinions or evaluations (subjective factors). Under this model, a highly-rated lawyer demonstrates excellence not only in objective data but also in client "positive reviews," which can be more persuasive than cold statistics. Indeed, this mechanism closely resembles platforms like Taobao or Dianping, where customer reviews (positive or negative) significantly influence consumer choices.

The chapter also examines a meaningful question: what are the three key factors that create a highly-rated lawyer? These are perception, experience, and engagement. Perception refers to how clients view each lawyer, reflected to some extent in client reviews. Quantitative data supporting this factor include a lawyer's average client rating and the percentage of clients who would recommend the lawyer. Experience denotes a lawyer's professional practice history, supported by quantifiable data combining three metrics: years in practice, number of certifications held, and number of professional awards received. Engagement reflects the degree of professional dedication and participation, supported by data on the number of articles published, speeches delivered, and peer recognitions received.

According to Avvo forum research, perception proves the most influential factor on overall lawyer ratings. This means that client experience, lawyer evaluations, and recommendations better reflect lawyer quality than actual practice experience or professional accolades. However, while perception is paramount, the other two factors also exert substantial influence. Statistical analysis reveals that all three factors are significantly correlated with lawyer quality, and each constitutes a necessary component of lawyer evaluation.

These findings yield important insights: lawyers with fewer awards, certifications, publications, or honors need not despair excessively, as focusing on client service and generating positive reviews offers a viable path to enhancing competitiveness. For less experienced practitioners, prioritizing respectful and appropriate client service to build positive perceptions can compensate for deficiencies in the experience factor.

Regarding how to generate positive client perceptions and improve the primary factor (reputation), this involves subjective evaluation. Client experience depends heavily on strong "client-lawyer personality fit." Research identifies specific behaviors lawyers can adopt to build positive client experiences and evaluations. Survey data from legal consumers on the Avvo forum indicates that three-fifths of respondents consider a lawyer's speed in returning calls and emails important in hiring decisions, making responsiveness a positive factor. One-third report that lawyers' body language matters: poor eye contact and inattention

alienate clients, leading to negative reviews and low ratings. One-quarter state that a lawyer's voice on the phone influences hiring decisions.

## Reflection on China's Lawyer Evaluation System

Within this chapter's theme, and based on research into lawyer evaluation systems, on March 27, 2019, the Ministry of Justice issued the "Notice on Expanding the Pilot Program for Lawyer Professional Competence Evaluation Systems and Assessment Mechanisms," explicitly stipulating the promotion of such systems across 31 provinces (autonomous regions and municipalities) and the Xinjiang Production and Construction Corps [CLI.4.330733]. From the lawyer-client relationship perspective, I offer the following reflections:

Regarding "quantifying value," the legal services industry often emphasizes the sacrifices behind a successful lawyer's practice—the need for lifelong learning and dedication. However, highlighting lawyers' hardships alone does not attract clients. When considering the issue from the client's perspective, what truly drives the decision to purchase legal services is whether the perceived expected value exceeds the cost. Therefore, the real motivator is the value clients feel. A basic strategy can involve three progressive steps: quantify value, emphasize differentiation, and provide added benefits. First, quantifying value requires lawyers' capabilities to be externalized as service products corresponding to specific client problems, giving clients an opportunity to perceive that value. This aligns closely with this chapter's discussion. Second, emphasizing differentiation reflects a common mindset: highlighting a lawyer's unique value in accomplishing client objectives. Finally, providing added benefits—common among lawyers in financial and corporate practice who work with various enterprise clients—involves the value of resource integration, whether through human resources services, financial and tax advisory, or investment-related services that can provide additional value to target clients.

## Data Privacy Protection in Legal Services

Connecting this chapter's discussion of Avvo forum's data acquisition and integration—particularly its high traceability and third-party platform data issues—I now examine data privacy protection, big data, and China's latest Personal Information Protection Law.

In his 1995 futurist work *Being Digital*, Nicholas Negroponte, hailed as a "digital prophet," wrote, "Computing is not about computers anymore. It is about living." Today's world finds itself on the "digital trajectory" he predicted, and the legal services industry is no exception. As we enjoy the dividends of the digital age and ride the wave of rapid digital development, have we considered its flip side—how to protect personal information security? How should data ownership rights be allocated? Can online legal services maintain the warmth of offline legal services?

### (1) Rights Defects in the Digital Context

Professor Mo Jihong has discussed rights defects in the digital realm. Drawing on ancient Greek "subject-object dualistic opposition" thinking, we artificially create a subject-object dichotomy, yet the world is fundamentally interconnected without such divisions. If we understand the world through the language of rights, we encounter challenges from modern digital technology. Because digital technology is mutable, the objects of rights become more uncertain. In the digital age, this breaks the "object myth," enabling winner-take-all scenarios. Those clever individuals who create new algorithms can control the world, making rights harmful to the less clever. Digital era development requires us to collectively recognize that digital technology needs morality, not just rights. How can clever people avoid manipulating others? This represents a crucial legal question in digital technology development. If those who wield legal weapons cannot provide genuine help to the vulnerable, rights cease to exist in our world. Therefore, we must change our thinking patterns, and legal professionals should lead the way on the path of morality.

### (2) Data Ownership Allocation

Information technology development involves two dimensions: AI for Law, where we empower law through information technology (such as smart document drafting), and Law for AI, where we create a sound legal environment for AI development to eliminate AI risks. As the digital economy develops and regions undergo digital transformation (such as the Guiyang Big Data Exchange), the absence of clear data ownership—who owns the data—constitutes the greatest obstacle to digital economic development. For instance, when Maimai scraped user data from Sina Weibo, Maimai argued: first, Sina Weibo user data is public; second, user data does not belong to Sina Weibo, so scraping is permissible. Sina Weibo contended that data generated on its platform, supported by its technology and capital, cannot be scraped by Maimai. Sina Weibo subsequently sued Maimai, creating the first major case of unfair competition triggered by big data.

Regarding data ownership allocation between platforms and enterprises, the logic should follow that whoever originates the data owns it. Thus, users as data originators should own the data, while platforms as data processors should have usufructuary rights to the data. This arrangement can achieve rights allocation between users and enterprises, reconcile conflicts of interest among different data enterprises, and ensure platforms' control and independent rights functions over data, allowing each party to remain in its proper place, observe its boundaries, and follow its path. However, future issues remain to be resolved, such as the establishment of data usufructuary rights and the principle of statutory property rights, the compatibility and exclusivity of data usufructuary rights, the dominance and disposability of data usufructuary rights, and the duration of data usufructuary rights and its determination basis.

### (3) Legal Professionalism in the Digital Era

The digital wave is irresistible, yet technological development should enable legal services to maintain their warmth without becoming lost in the debate between legal professionalism and legal commercialism. Three questions warrant consideration and further research under professionalism. First, regarding lawyer and law firm digitalization: should there be standards rather than simply chasing trends? How should such standards be formulated, and by whom? Second, any innovative undertaking requires substantial investment of human, financial, and material resources. Does the legal profession possess such conditions, and how can it obtain them? China has over 520,000 lawyers and 37,000 law firms, with a GDP of approximately 200 billion RMB in 2020. However, 96% of these firms are small and medium-sized, with fewer than 50 lawyers. We must consider who will use the fruits of digitalization and whether these small firms have the conditions for digital transformation. Third, legal services constitute a warm-hearted profession with an indelible shadow of traditional legal professionalism. After digital transformation, how can legal services avoid losing this perceptible warmth, or even become warmer? This requires consideration amid the wave of legal commercialism.

### (4) Privacy Policy Issues under the Personal Information Protection Law

Based on the chapter's discussion, data protection on third-party platforms like Avvo is a significant issue. Analogizing lawyers to commodities and platforms like Avvo to Taobao or Dianping, their privacy policies regarding various data and information warrant attention. Formulating privacy policies represents an important tool for self-regulation by online service providers and a crucial method for ensuring compliance with personal information protection legislation. Privacy policies approved by users establish a contractual relationship between users and service providers, granting authorization for collecting and using personal information. Changes to such privacy policies constitute contractual modifications requiring user consent; otherwise, they are not binding on users. Privacy policies not agreed to by users are purely corporate self-regulation rules that cannot establish contractual relationships and do not grant authorization for collecting and using personal information. Changes to such policies need not obtain user consent, and modified policies remain non-binding on users.

### (5) Data Protection Laws and Key Protection Points

**1. Participants and Key Data Protection Points** Throughout the entire data lifecycle—collection, storage, transmission, use, analysis, sharing, trading, disclosure, and destruction—parties involved in data security can be categorized into three types: data subjects, data controllers, and data regulators.

**(1) Data Subject:** Refers to the natural person or organization to whom the data pertains, typically users or customers of information services.

**(2) Data Controller:** Refers to natural persons or organizations that collect, transmit, store, use, or otherwise process data, typically individuals or institutions providing information services. Data controllers may be one or more natural persons, companies, associations, or third-party data controllers introduced through service outsourcing (generally called data intermediaries). Typically, data controllers have control over data and can determine the purposes and means of data processing.

**(3) Data Regulator:** Refers to institutions that formulate data processing policies and security rules, supervise data processing security issues, accept complaints and reports, and penalize improper conduct, typically government agencies or institutions with government backgrounds.

Based on a survey of national laws and regulations, the purpose of data protection legislation in all countries revolves around these three participants, seeking to clearly define their boundaries of responsibility, corresponding rights and obligations, and relevant codes of conduct.

## 2. Data Regulator Level

**2.1 Scope of Data Protection** Data protection has scope; not all behaviors of all data controllers or all collected data require protection. Protection should be limited to the regulatory jurisdiction, targeting data controllers, behaviors, and data that require regulation. Therefore, when formulating data protection laws, regulators must first define the regulatory jurisdiction, protected data, regulated objects, and regulated behaviors.

**(1) Regulatory Jurisdiction:** Refers to the scope of data covered by laws and regulations, particularly whether data centers established overseas fall under national regulatory oversight. Different countries and regions have varying provisions. The United States, Australia, and China currently limit jurisdiction to national territory. However, the EU, Russia, Singapore, and others have stricter regulations. Russia stipulates that its data protection laws are not territorially limited but apply to all data processing occurring in Russia, including all collection, sale, and use of Russian citizens' data, regardless of whether data centers are established or located within Russia. For cross-border data flows, if Russian citizens are parties to the corresponding data transfer agreement, the data falls within Russian data protection regulatory jurisdiction.

**(2) Protected Data:** Generally, data requiring regulation divides into two categories: personally identifiable information and personal privacy/sensitive data. Identifying information enables direct identification and location of individuals, such as names, ID numbers, bank card numbers, and home addresses. Personal privacy and sensitive data may not directly identify individuals but could locate them through correlation and comprehensive analysis, such as health information, educational history, and credit records. Countries define personal privacy and sensitive data differently, resulting in varying scopes of protected

data. China delineates specific personal information protection scopes in departmental regulations, while Russia and Singapore consider all information related to individuals as personal privacy and sensitive data within protection scope. Additionally, exemption clauses exclude certain data from regulation, such as Singapore's provisions that business contact information, personal data existing for 100 years, and personal data of deceased individuals beyond ten years fall outside protection scope.

**(3) Regulated Objects:** Generally, all data controllers involved in data collection, sale, storage, processing, and utilization are regulated objects. However, countries have formulated exemption clauses based on national conditions, such as Singapore's provisions exempting personal citizen behavior, necessary employee conduct, partial behavior by government, press, and research institutions, and certain data intermediaries with explicit certifications or written contracts.

**(4) Regulated Behaviors:** Currently, the United States, EU, China, Russia, Singapore, and other countries propose regulating the entire data lifecycle, including collection, recording, organization, accumulation, storage, alteration (updating, modification), retrieval, recovery, use, transfer (dissemination, providing access, etc.), de-identification, deletion, and destruction. However, countries have also formulated exemption clauses, such as Russia's provisions that processing personal data exclusively for personal and family needs (provided they do not infringe on data subjects' rights), processing state secret data, and providing relevant data to Russian courts by competent authorities under court legislation constitute exceptions.

**2.2 Regulatory Authorities and Powers** To ensure implementation of data security laws, regulators must establish institutions and personnel with corresponding authority. Singapore's primary legal basis for data protection is the Personal Data Protection Act, enforced by a specially established Personal Data Protection Commission responsible for formulation and implementation. China has not yet established a dedicated national regulatory agency to determine compliance with personal information usage. Instead, various administrative departments regulate within their purview: the Ministry of Industry and Information Technology oversees personal information in telecommunications and internet industries, the National Health and Family Planning Commission supervises medical records and resident health information, and the State Administration for Industry and Commerce regulates consumer personal information.

**3. Data Subject Level** The data subject level focuses on defining data protection points around data subjects, namely users' rights when using information services. These include the right to be informed before data collection and processing (right to know), the right to authorize personal data collection and processing (right to authorize), the right to access, inquire about, and correct personal information, the right to stop collection and delete personal information (right to cessation and deletion), the right to complain, and other related

rights. National data protection laws generally stipulate data subject rights, but to varying degrees. Singapore requires data controllers to obtain data subjects' consent before collecting personal data but does not specify particular notification forms.

**4. Data Controller Level** The data controller level focuses on defining data protection points around obligations throughout the data lifecycle, including obligations to cooperate with data subjects in exercising their rights, ensure data security, obtain consent from data regulators before collection, report data collection and utilization to regulators, notify regulators of abnormal incidents, and apply to regulators before cross-border data transfer or storage. Regarding the obligation to report data collection and utilization, China's laws do not explicitly require data controllers to apply for or register with designated government authorities before processing data. Russia requires data controllers to report to their supervisory authorities before processing personal data (via paper or electronic submission), including the controller's name and address, purposes and types of personal data processing, data subject categories, data protection officer's name and contact information, and database addresses containing Russian citizens' data. Authorities register the controller's information within 30 days and publish it on relevant websites. Singapore requires data controllers to notify authorities in general, without specifying forms, but particularly notes that for data usage purposes unknown to users, controllers must provide all relevant information.

The above discussion covers five aspects: rights defects in digital context, data ownership allocation, legal professionalism in the digital era, privacy policy issues under the Personal Information Protection Law, and data protection laws and key points, with brief mentions of select foreign data protection regulations. Future research will continue to explore issues around "legal services and judicial management technology," delving deeper into data protection, big data, and legal technology issues in legal service institutions. I believe many more aspects of "law-technology" merit further investigation.

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