

Generative AI in Legal Tech: A Framework for Effective Solution Evaluation

By Miguel Villalobos

April 28, 2025

Generative AI has rapidly become central to legal technology discussions, with use cases spanning e-discovery, investigations, compliance and more. However, navigating the myriad solutions can be daunting. Legal teams can benefit from a robust framework to effectively assess the potential and pitfalls of gen AI before full-scale adoption.

Here's a structured approach that will enable organizations to establish the key use cases for gen AI and then select the right tools.

Identifying Key Use Cases—Solution Value vs. Organizational Scrutiny

When evaluating gen AI solutions, leaders should consider the balance between the tool's potential value and the scrutiny the use case will attract. Use cases that are considered "high scrutiny" are those where accuracy and precision are paramount, and outcomes can significantly affect legal results.

These high-stakes use cases might require—and justify—the adoption of more sophisticated, costly tools that can deliver the required accuracy and efficiency. But such investments could be worth it for high-value use cases, where the use of gen AI can significantly impact outcomes. Examples of high-value, high-scrutiny use cases include data categorization and culling, where accuracy and precision are critical to the final result, and the effi-



Credit: Thanadon88/Adobe Stock

ciency potential of gen AI could be transformative to the organization.

Conversely, lower scrutiny scenarios, such as internal preliminary document assessments or routine document organization tasks, may permit greater flexibility in experimenting with innovative or more cost-effective AI solutions, even if their immediate value might be incremental. Understanding this interplay helps legal teams prioritize investment, resource allocation, and risk management effectively.

Key Assessment Criteria

Once use cases have been identified, the following framework can be used to assess specific gen AI tools:

1. User-Friendly Interface (UI): Evaluate the intuitiveness and ease of navigation of the platform's interface. A user-friendly AI platform is not merely nice to have: it lowers the risk of sunk cost by accelerating adoption, minimizing training overhead, and ensuring consistent use by legal teams.

2. Workflow Adaptability: Assess how easily the tool integrates into existing document review workflows and platforms, such as Relativity. Flexible tagging schemes and seamless workflow integrations are vital for efficiency gains. Evaluate the landscape of any proprietary technology partnerships that may hamper effective integration, now or in the future.

3. Cost: Analyze not only licensing fees but also indirect costs such as training data requirements, ongoing model tuning, and infrastructure maintenance. Transparent and predictable costs are essential for budgeting and long-term planning.

4. Market Momentum: While not a direct measure of technological performance, leaders should evaluate the vendor's market momentum—including widespread adoption, client testimonials and ongoing vendor innovation. These can be critical indicators of reliability and longevity.

5. Accuracy: Precision, recall and factual consistency metrics should be rigorously evaluated to ensure the AI solution performs as advertised. Depending on the use case, teams should ensure the technology accurately generates text, extracts information, identifies, categorizes, translates and summarizes relevant documents, to ensure outputs align with legal standards, minimize risk, and support reliable decision-making. Obtaining accuracy metrics is relatively straightforward for coding and some extraction use cases by relying on standard TAR validation methods. Validating other outputs benefits from using text focused techniques such as calculating ROUGE or BLEU scores.

6. Customization and Prompt Engineering: The flexibility to tailor prompts and AI behavior to the nuances of a legal case greatly impacts effectiveness. Legal teams should understand their own potential customization needs based on the use case and ensure the tool can align its outputs with tailored legal strategies and case requirements.

7. Performance at Scale: E-discovery and other legal workflows often involve vast amounts of data. Tools should be evaluated for efficiency and reliability when handling large-scale document sets.

8. Data Privacy and Geographic Deployment: Legal teams should evaluate the platform's compliance with data privacy standards and information security protocols and understand the type of Large Language Models (LLMs) employed. Vendors should be able to demonstrate that they can meet any applicable jurisdiction-specific data privacy regulations.

Positioning for the Future

As generative AI reshapes the legal technology landscape, a structured evaluation framework becomes essential. By aligning solution selection with organizational priorities—balancing value creation against organizational scrutiny—legal teams can make more confident, strategic investments. Thoughtful assessments ensure that selected technology works harmoniously with an organization's existing technology, operations and regulatory constraints. The firms that succeed amid today's rapid innovation will be those that match their ambition with rigor, adopting gen AI tools that elevate both performance and compliance across legal workflows.

Miguel Villalobos is senior director, AI and e-discovery, at Integreon, a provider of legal and business outsourced services to corporations and law firms worldwide.