ISSUE N.4 | DICEMBRE 2024

Diritto Intelligente

A JOURNAL ON AI-RELATED EU LAWS, CASES, AND OPINIONS BY DLA PIPER'S ITALIAN INTELLECTUAL PROPERTY AND TECHNOLOGY GROUP.

In this issue

DLA PIPER

Contents

Editorial

Embracing the era of AI agents: What to expect in 2025?

As we approach the close of another transformative year in the world of artificial intelligence, it is with great enthusiasm that we present the December issue of *Diritto Intelligente*. This month, we turn our focus to what is arguably the most significant innovation of recent times: AI agents. These autonomous systems are not just a leap forward in technology; they represent a paradigm shift in how we interact with machines and, consequently, how we must navigate the evolving legal landscape.

AI agents have moved beyond simple text-based interactions to become capable of performing complex tasks on our computers, mimicking human actions such as moving cursors, typing, and managing applications.

In this issue, we delve into the multifaceted legal challenges posed by AI agents. We explore the privacy and security concerns that arise when these systems require deep access to personal and corporate data, emphasizing the need for robust safeguards and a balanced legal framework. We underscore the importance of proactive policymaking to ensure that the integration of AI agents is both beneficial and responsible.

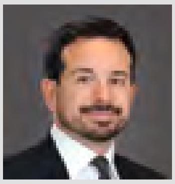
Building on this theme, we provide an insightful analysis of the recently published first draft of the General-Purpose AI Code of Practice. This draft aims to detail the AI Act rules for providers of general-purpose AI models, addressing key considerations such as transparency, systemic risk assessment, and risk mitigation strategies. As the final version is anticipated by May 2025, we offer a timely perspective on the regulatory efforts shaping the future of AI governance.

We also examine the reinforced US strategy on artificial intelligence, highlighting the National Security Memorandum that outlines the country's approach to maintaining leadership in AI while addressing national security concerns. This perspective offers a global context to the advancements and regulatory responses surrounding AI agents.

Lastly, we discuss a significant move by a major publishing house to exclude its books from being used to train AI technologies. This development signals a growing awareness and assertiveness among content creators regarding the use of their intellectual property in AI training, raising important questions about rights and compensation. This decision not only challenges current practices in AI development but also prompts us to consider how intellectual property laws will adapt to these technological advancements. Will other publishers follow suit, leading to a more restrictive data environment for AI training? How will this impact the evolution of AI technologies that rely heavily on vast amounts of data? These questions highlight the pressing need for a dialogue between content creators, AI developers, and legal professionals to find a balanced approach that protects rights while fostering innovation.

As we look ahead to 2025, we anticipate that these discussions will intensify, and new legal challenges will emerge. The rapid advancement of AI agents will undoubtedly test the boundaries of existing laws and regulations. It is imperative for us, as a legal community, to stay engaged, informed, and collaborative in addressing these issues.

It has been an amazing year in the world of innovation, the whole Italian Intellectual Property and Technology team at DLA Piper wishes all our readers a joyous holiday season and a fantastic start to 2025. We look forward to working together in the coming year to explore the new legal challenges that await us, ensuring that the integration of AI into our society is both innovative and responsible.



Giulio Coraggio
Location Head of the Italian Intellectual Property
and Technology Department at DLA Piper

AI Agents: What legal implications of autonomous artificial intelligence?

Author: *Giulio Coraggio*

The world of **artificial intelligence (AI)** is rapidly evolving, and a new milestone has been reached with the launch of advanced **AI agents** that bring significant **legal** considerations.

These agents are poised to transform our interaction with technology by not only responding to text commands but also performing actions on our computers as a human would: moving the cursor, typing, clicking, and reading the screen. Imagine a future where your computer doesn't just answer your queries but actively assists you by navigating windows, filling out forms, and managing your tasks. While this represents a significant advancement in **AI technology**, it also brings forth a host of **legal challenges** that must be carefully addressed.

Beyond text: What do advanced AI agents do?

Traditional **AI chatbots** have been confined to responding within the boundaries of text. However, the new generation of advanced **AI agents** breaks free from this limitation. Developers can now program AI that interacts directly with the computer environment, automating repetitive and mundane tasks. Although these systems are still in their infancy—prone to errors and operating at slower speeds—they signal the beginning of a shift towards **AI agents** handling more complex activities autonomously.

For instance, an **AI agent** could gather information from your computer and complete forms without human intervention. This might seem trivial, but the implications are far-reaching. Such capabilities could revolutionize productivity by offloading routine tasks from humans to **AI agents**, allowing individuals to focus on more strategic and creative endeavors.

Several tech giants and startups are investing in similar **AI agent technologies**. What sets these agents apart is their ability to act beyond text, directly interfacing with computer systems to manage intricate projects with greater autonomy.

Privacy and security concerns of AI agents

The advent of **AI agents** capable of operating our computers raises significant privacy and security issues. To function effectively, these agents require direct access to our devices, which poses risks of data breaches and unauthorized data transmission. There have been instances where companies delayed launching similar functionalities due to security concerns. Ensuring that user data is protected and used securely is paramount in the **legal landscape** of AI.

Moreover, granting such deep access to **AI agents** could inadvertently expose sensitive personal or corporate information. Without robust security measures, there is a heightened risk of malicious exploitation by cybercriminals who might hijack these agents for nefarious purposes.

This pivotal moment in the future of **artificial intelligence** is happening just days before the European Data Protection Board's event dedicated to AI models. Hopefully, data protection authorities will understand that generative **artificial intelligence**, including advanced **AI agents**, is the future of our economy. Solutions need to be found to adequately balance the protection of individuals with the exploitation of such technologies within the **legal framework**.

The liability cannot rest solely on providers of generative **AI**. Potential misuses are performed by deployers who may not have a clear understanding of the **legal limits** within which such technologies should be used.

Potential for misuse of AI agents

The power of advanced **AI agents** also opens the door to potential misuse. Autonomous navigation capabilities could be exploited for activities like spamming, phishing, or generating large-scale AI-created content that floods digital spaces. The ease with which these agents can perform tasks might enable the rapid dissemination of misinformation or the creation of fraudulent schemes.

This scenario underscores the necessity for clear, responsible management that needs to be implemented by setting out an internal **AI governance framework**, leading to precise internal rules and technical guardrails in **AI usage**. Developers and companies must implement safeguards to prevent abuse, such as strict authentication protocols, usage limits, and monitoring systems to detect and halt suspicious activities, all within **legal** boundaries.

Legal challenges to address

With the emergence of such potent **AI agents**, companies adopting these technologies must address several **legal challenges** that demand attention:

- **Data Protection and Privacy**: How can we ensure that **AI agents** do not access or transmit personal data outside the control of the company and without the relevant legal basis? Compliance with data protection laws like the GDPR becomes even more critical in the realm of **AI**.
- **Liability Issues**: In cases where an **AI agent** makes an error or causes harm, determining liability becomes complex. Is it the developer, the user, or the **AI agent** itself? This poses significant **legal questions** that need clear answers.
- **Intellectual Property Rights**: **AI agents** that create content or gather data from various sources may infringe on intellectual property rights, leading to **legal disputes**.
- **Regulatory Compliance**: Existing laws may not adequately cover the capabilities of advanced **AI agents**. There's a pressing need for updated regulations that address these new **AI technologies** within the current **legal system**.

Looking beyond the "Text box"

The shift from text-based AI to agents that can interact with our computers is a game-changer. While current versions may be imperfect, continual improvements will enable these **AI agents** to handle increasingly complex tasks. We can foresee a future where **AI agents** manage appointments, fill out forms, respond to emails, and even curate personalized news briefings without any manual input.

However, embracing this future necessitates confronting the accompanying **legal and ethical implications**. Security, privacy, and ethical use are not just technical challenges but **legal ones** that require collaboration between technologists, legal experts, policymakers, and society at large.

Conclusion

The advent of advanced **AI agents** heralds an exciting new chapter in **AI technology**, offering unprecedented convenience and efficiency. Yet, it also brings forth significant **legal challenges** that cannot be overlooked. Addressing these issues is crucial to harnessing the full potential of **AI agents** while safeguarding users' rights and maintaining public trust.

As we stand on the cusp of this technological revolution in **artificial intelligence**, it is imperative to engage in open dialogue and proactive policymaking. By doing so, we can ensure that the integration of **AI agents** into our daily lives is both beneficial and responsible within the **legal framework**.

As mentioned above, whatever AI solution a company wants to adopt, a crucial step in the adoption relates to the creation of an AI governance framework. Feel free to contact us if you have any questions on the topic, and try our PRISCA AI Compliance tool described HERE.

First draft of the general-purpose AI code of practice published

Author: *Roxana Smeria*

A group of independent experts presented the first draft of the General-Purpose AI Code of Practice, that aims to detail the AI Act rules for providers of general-purpose AI models and general-purpose AI models with systemic risks.

This first draft of the Code addresses key considerations for providers of general-purpose AI models and for providers of general-purpose AI models with systemic risk. Although the first draft is light in detail, this approach aims to provide stakeholders with a clear sense of direction of the final Code's potential form and content. Key areas of focus include transparency, a taxonomy of systemic risks, robust risk assessments, and stringent risk mitigation strategies, encompassing both technical and governance measures. The final version of the Code should be ready by 1 May 2025.

Transparency

A primary focus of the draft is ensuring transparency in the development and use of AI models. Providers will be required to draw up and keep up-to-date the technical documentation of the model listed in the Code (such as intended tasks and type and nature of AI systems in which it can be integrated, acceptable use policies, interaction of the model with external hardware or software) for the purpose of providing it, upon request, to the AI Office, the national competent authorities and to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems. Providers are also encouraged to make elements of this information publicly available to promote transparency and trust.

With specific reference to the Acceptable Use Policy, providers should commit to sharing with the downstream providers all the necessary information related to their general-purpose AI model to enable downstream providers to comply with existing regulations applicable to the task or use case their AI system is intended to be used for. The Code provides also a list of elements that shall be included in the Acceptable Use Policy, such as (i) the scope defining who the policy applies to and what resources it covers; (ii) primary intended uses and users; (iii) unacceptable uses, detailing forbidden actions; and (iv) security measures containing a description of the security protocols that the users of the general purpose AI systems must follow.

Taxonomy of systemic risks

The draft introduces a comprehensive taxonomy of systemic risks associated with general-purpose AI models. Providers shall commit to draw from the elements of this taxonomy of systemic risks as a basis for their systemic risk assessment and mitigation.

According to the draft, signatories shall treat the following as systemic risks:

- **Cyber offence**: Risks related to offensive cyber capabilities such as vulnerability discovery or exploitation.
- **Chemical, biological, radiological, and nuclear risks**: Dual-use science risks enabling chemical, biological, radiological, and nuclear weapons attacks via, among other things, weapons development, design, acquisition, and use.
- **Loss of Control**: Issues related to the inability to control powerful autonomous general-purpose AI models.
- **Automated use of models for AI Research and Development**: This could greatly increase the pace of AI development, potentially leading to unpredictable developments of general-purpose AI models with systemic risk.

- **Persuasion and manipulation**: The facilitation of large-scale persuasion and manipulation, as well as large-scale disinformation or misinformation with risks to democratic values and human rights, such as election interference, loss of trust in the media, and homogenisation or oversimplification of knowledge.
- **Large-scale discrimination**: Large-scale illegal discrimination of individuals, communities, or societies

When determining a systemic risk, signatories shall also consider the nature (such as intent, novelty, velocity at which the risk materializes) and the source.

Risk assessment

Providers of general-purpose AI models are required to adopt rigorous risk assessment methodologies to ensure the identification, evaluation, and mitigation of systemic risks throughout the model's lifecycle. Key requirements include:

- **Robust Methodologies**: Providers must employ sophisticated and reliable risk analysis techniques to identify potential pathways through which AI models might pose systemic risks. This includes estimating the likelihood and severity of such risks materializing, ensuring proactive management.
- **Mapping Systemic Risk Indicators**: An essential part of the process is identifying and documenting the capabilities or tendencies of AI models that may be linked to systemic risks. These risk sources must be mapped to specific indicators that can serve as early warning signs for emerging threats.
- **Severity Tiers**: Identified risks must be classified into severity tiers, with clear distinctions between manageable risks and those deemed intolerable. This ensures that critical threats are promptly addressed with appropriate safeguards.
- **Risk Forecasting**: Providers are also expected to anticipate when systemic risks are likely to arise, offering best-effort estimates based on the model's development trajectory and application environment. This forward-looking approach allows for early interventions to prevent risks from escalating.

In addition to these specific measures, providers are obligated to maintain a continuous risk assessment process across all stages of a model's lifecycle. This involves regularly gathering and analyzing evidence to monitor risk indicators and the effectiveness of mitigation strategies. Assessments are required before and after implementing risk mitigations, ensuring that the measures taken remain robust and relevant in the face of evolving challenges.

Technical and governance risk mitigation

The draft emphasizes the need for both technical and governance-based measures to mitigate systemic risks.

Technically, providers shall link systemic risk indicators to proportional safety and security measures to keep risks below intolerable levels and minimize them further. Providers, therefore shall put in place the following:

- **Safety Mitigations**: including adjust models' behavior, safeguard deployment systems, or provide countermeasures to mitigate systemic risks.
- **Security Mitigations**: meaning that providers shall protect unreleased model weights and assets during and after development, using measures like access control, monitoring, and red-teaming.
- **Limitations**: Providers shall document in their safety and security framework the limitations in existing safety and security mitigations to identify gaps in mitigation measures.
- **Adequacy Assessment**: Providers shall regularly evaluate in the safety and security framework the effectiveness of mappings between risks and mitigations.

Moreover, providers must create detailed safety and security reports at key development stages, documenting risk and mitigation assessments and establish criteria to halt or continue model development based on safety and security reports results.

As per the governance mitigation, providers shall ensure accountability for systemic risks at executive and board levels, allocating resources and establishing oversight committees. Moreover, annual evaluations shall be conducted to assess adherence to the safety and security framework and its relevance to evolving risks and practices.

Providers shall also enable independent evaluations of systemic risks and mitigations throughout the model lifecycle. Such independent expert risk and mitigation assessment may involve independent testing of model capabilities, reviews of evidence collected, systemic risks, and the adequacy of mitigations. Before deployment providers shall facilitate external testing and review, after deployment, providers shall support ongoing assessments to address emerging risks.

Finally, providers shall:

- establish processes to identify, document, and report serious incidents or near-misses to the AI Office and define corrective measures;
- implement and inform employees about secure channels for reporting violations, with appropriate safeguards;
- notify the AI Office about models meeting systemic risk thresholds, updates to the safety and security framework and safety and security report, and emerging systemic risks;
- maintain detailed records throughout the AI model lifecycle to demonstrate compliance with risk mitigation standards and facilitate AI Office requests;
- publish safety and security frameworks and safety and security reports, balancing societal benefit with the need to protect sensitive information.

Conclusions

The draft General-Purpose AI Code of Practice lays the groundwork for a regulatory framework that balances innovation with safety and societal accountability.
The upcoming stakeholder consultations will refine the draft, integrating feedback to create a more comprehensive and effective Code.

The final version, due in May 2025, will serve as a vital tool in ensuring that the deployment of general-purpose AI models adheres to ethical and legal standards, minimizing systemic risks while fostering trust in AI technology.

US strategy on artificial intelligence is reinforced

Author: *Giacomo Lusardi*

Ahead of the Nov. 5, 2024 presidential election, the US administration adopted on **Oct. 24** a National Security Memorandum (Memorandum) titled "*Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence*," the adoption of which was mandated by the **US Executive Order of Oct. 30, 2023** on the safe, reliable, and accountable development and use of artificial intelligence (AI).

The meaty and articulate Memorandum is the most up-to-date document on the U.S. national security strategy on AI. Also on Oct. 24, the White House released a **complementary document** to the Memorandum, titled "*Framework to Advance AI Governance and Risk Management in National Security*."

Let's briefly explore the **key themes** of the Memorandum and its **intended audience**, as well as its **potential fate** in the future.

Frontier artificial intelligence

First, unlike the 2023 Executive Order, the Memorandum primarily focuses on **generative AI**, a wave that began with the release of OpenAI's **ChatGPT** in **2022**. Generative AI models, such as those powering OpenAI's ChatGPT, Anthropic's Claude or Google's Gemini, are an **evolution** from previous *deep learning* models because they are adaptable across a **wider range of** applications. In contrast, the **previous generation** of AI, primarily based on supervised learning (often called supervised machine learning), was more tailored to **specific applications** and, as a result, was generally more predictable and posed lower risks.

Evoking imagery dear to American culture, the Memorandum names these models "**frontier models**," defining them as "*general-purpose AI systems near the cutting-edge of performance, as measured by widely accepted publicly available benchmarks, or similar assessments of reasoning, science, and overall capabilities*." A definition that seems to echo that of the European AI Act, which went into effect last **August 24**, in which **general-purpose** AI models, or, in English, "*General Purpose Artificial Intelligence*," are AI models characterized by **significant generality** capable of competently performing a **wide range of distinct tasks**.

Who the memorandum addresses, key themes and the role of China

The Memorandum establishes **US national security policies** with respect to frontier AI, assigning specific **responsibilities** to various federal agencies. On the one hand, it highlights that the government should support and nurture **the leadership of** the U.S. AI industry and, on the other hand, what the government itself should expect from the **private sector** to succeed in achieving national security goals.

In more detail, the document sets out a series of measures to ensure that the United States maintains its **position of primacy** in the global AI ecosystem. Essential in this regard **are talent attraction** and the government's ability to provide **appropriate security** and **protection guidelines** to AI developers and users, helping to mitigate the **risks that** AI systems can pose. According to the Memorandum, **security and reliability of** AI are crucial aspects of accelerating the adoption of AI systems, and **the absence of** clear **guidelines** can be **an obstacle**. In addition, the Memorandum designates **the** "**AI Safety Institute**" (**AISI**) at the Department of Commerce as the primary **point of contact for** AI companies in the governmental sphere in relation to the **evaluation** and **testing** activities of AI systems.

Another crucial aspect touched upon in the Memorandum is the need for large-scale expansion of **IT infrastructure** and **data centers** to fuel the growth of the AI industry very quickly.

As for **governance**, the document charges national security agencies with several **tasks**. Nearly all agencies will be required to designate a "***Chief AI Officer***" (**CAIO**) and a National Security **Coordination Group** for AI will be created composed of the CAIOs of the major agencies. The Memorandum also encourages U.S. **cooperation** with **international** partners and **institutions**, such as the G7, OECD and the United Nations, to **promote** international AI **governance**.

What about **China**? It is **never mentioned** directly in the Memorandum (it is referred to, generically, as "competitors"), but it is undoubtedly the **main competitor** to the United States for global AI leadership. Part of the document describes how the U.S. intends to surpass competitors in this race and emphasizes how partners and allies play a **central role**. On the point, last **Oct. 28** the US administration also published the long-awaited *final rule* to **limit U.S. investment in China**. The control regime, which will go into effect **on 2 January 2025**, will affect all U.S. companies and citizens who invest in Chinese companies operating in the fields of AI development, semiconductors and microelectronics, and quantum information technologies.

The Memorandum presents an ambitious and detailed vision of **AI's role** in U.S. **national security**, we will see how the new administration will address the issue.

A major publishing house explicitly excludes its books and reprints from being used for AI training

Author: *Chiara D’Onofrio*

One of the major publishers worldwide has been reported to have added an explicit reference to artificial intelligence (AI) in its newly published books and reprints’ copyright pages, stating that no part of these works can used or reproduced in any manner “for the purposes of training artificial intelligence technologies or systems”. This is the first example of a publishing house taking action against the exploitation of published paper and digital works to train AI technologies, including large language models (LLMs).

Copyright notices are used by publishers to assert their rights and those of their authors over printed and digital books. They are also used to inform the reader of what can and cannot be legitimately done with the work. In any case, where copyright disclaimers are not used, existing copyright protections still apply.

By adding a reference to AI system training in its copyright notice, the publishing house is effectively excluding its works from being used to develop chatbots and other AI digital tools, which has allegedly been done in the past by using published (and pirated) books without the consent and authorization of rightsholders.

Further, in the newly drafted copyright notice to be added to books, the publisher expressly reserves its works from “the text and data mining exception” in accordance with Article 4(3) of Directive (EU) 2019/790 on copyright and related rights in the Digital Single Market (“**Copyright Directive**”).

According to Article 2 of the Copyright Directive, text and data mining is “*any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations*”. As set by Article 4 of the Copyright Directive, exceptions or limitations to exclusive rights on copyright-protected works, that are legitimately accessed, can be provided to allow *text and data mining* activities to be carried out. Paragraph 3 of Article 4 of the Copyright Directive specifies that such exceptions or limitations are applicable only on the condition that rightsholders have not expressly reserved the use of their works in an appropriate manner, such as machine-readable means in the case of content made publicly available online.

The newly drafted copyright notice of one of the “Big Five” publishing houses should work in a similar way to exclusion protocols contained in ‘robots.txt’ files, which are being used by websites to exclude their content from scraping activities by bots and AI technologies. This can be interpreted as a first step in the publishing industry to adopt clearer and explicit statements by publishers on reserving training, text and data mining rights in relation to their published works.

The decision by one of the major publishing companies to rewrite its copyright notice further highlights the ongoing tension between content creators and the AI world. A growing number of publishers, authors, and players in the sector are requesting stronger and more defined protection of their exclusive rights and are taking a defensive approach towards AI technologies and their training and output production in generative uses. For example, it has been reported that another prominent publisher recently adopted explicit measures prohibiting freelancer collaborators working on its authors’ books from copying any of the information and text contained by books into AI systems and programs for the purposes of editing, checking, extraction, or any other related purpose.

In addition to the measures adopted by publishers, authors and their representative organizations are calling for changes in publishing contracts with appropriate safeguards for creators. Agreements with publishers should ensure that authors’ consent is obtained before publishers use or allow the use of the works to train AI systems and, more generally, before granting access to the work to an AI system, for example, to produce AI-generated translations, audiobooks, and cover art of copyright-protected books.

In addition, several representatives of the creative industries have been advocating for the introduction of a more comprehensive legal framework. This framework should provide for adequate and transparent licensing provisions to ensure that creators and rightsholders are adequately paid for the use of their works, including in the context of uses made by AI systems. To this extent, new machine-readable text and data mining licenses are being devised to provide legitimate access, in some instances with payment, to machines that automatically scrape copyright-protected content.

Ultimately, rightsholders, both authors and publishers, are increasingly more interested in retaining meaningful control over how and to what extent their works interact with AI systems while being fairly and rightfully compensated for any exploitation of their works.

Legal design tricks

LITTLE TIPS TO USE LEGAL DESIGN IN YOUR DAILY ACTIVITIES

Trick #3: EMPATHIZE – The first step towards design thinking.

Author: *Deborah Paracchini*

What does "empathize" mean?

Empathize means making legal solutions more human-centered by truly understanding users' experiences and pain points. By empathizing with our audience, we can better tailor our solutions to address their unique needs.

How to empathize with users?

1. **Listen Actively and Observe Behavior –** Engage fully, focusing on both spoken words and non-verbal cues.
2. **Ask Open-Ended Questions –** Encourage users to share in-depth insights and personal experiences though surveys and interviews.
3. **Map User Journeys –** Visualize key touchpoints and pain points in users' interactions with legal processes.
4. **Reflect and Validate –** Develop solutions that truly align with user needs, reduce friction, and enhance satisfaction.

Why empathy?

Empathy is the foundation of user-centered legal solutions. It allows us to step into the shoes of our users, gaining insight into their needs, concerns, and frustrations.

What to do in practice?

Identify "personas" by creating detailed profiles of your model user. This involves exploring elements like name, age, family, profession, hobbies, and habits. Through this, you gain a clear picture of the *persona*'s needs, behaviors, and preferences.

Did you know?

Personas aren't always based on real individuals – they can be creatively imagined by the design thinker to represent typical users.

Find new tips every month on Diritto Intelligente or check our monthly posts at dirittoaldigitale.com

Legal tech bytes

EXPERT INSIGHTS ON THE LATEST TRENDS AND INNOVATIONS

Fine-tuning vs RAG: Tips for an effective AI implementation in Legal activities

Author: *Tommaso Ricci*

The legal technology landscape is undergoing significant transformation as organizations aim to leverage the capabilities of Large Language Models (LLMs). Central to this evolution is the important decision between two methodologies: Retrieval-Augmented Generation (RAG) and Fine-tuning. Understanding these approaches has become crucial for legal professionals who are increasingly under pressure to provide more efficient, accurate, and scalable solutions.

Understanding the fundamentals

Large Language Models have revolutionized how we interact with artificial intelligence, leveraging massive pre-training on diverse datasets to generate text, answer questions, and perform complex language tasks. However, their broad knowledge comes with limitations, particularly in specialized domains like law where precision and context are paramount.

Retrieval-Augmented Generation (RAG) addresses these limitations by combining an LLM's language understanding capabilities with domain-specific knowledge retrieval. This approach allows legal organizations to augment their LLMs with precise legal documentation, precedents, and internal knowledge bases. Rather than relying solely on the model's pre-trained knowledge, RAG systems can pull relevant information from verified sources before generating responses.

Fine-tuning, in contrast, takes a different path by modifying the LLM itself through additional training on specialized legal datasets. While this approach can create more specialized models, it often comes with significant tradeoffs in terms of model versatility and maintenance requirements.

The choice between RAG and fine-tuning also has significant cost implications. While RAG systems require investment in knowledge base maintenance and retrieval infrastructure, these costs are often more predictable and scalable than the computational resources required for regular fine-tuning of large models. For smaller legal organizations, RAG often presents a more accessible entry point into AI adoption.

Below is a high level table comparing the main differences between the two methodologies:

ASPECT	RAG	FINE-TUNING
Knowledge Updates	Real-time updates possible through knowledge base modifications; no retraining needed	Requires complete model retraining to incorporate new knowledge
Cost Structure	Higher operational costs (storage, retrieval), lower training costs	High initial training costs, lower operational costs
Accuracy and Reliability	High accuracy with proper retrieval; clear source attribution	Can be more accurate for specific tasks but risks "hallucination" when facing novel scenarios
Scalability	Easily scalable across different legal domains by updating knowledge bases	Requires separate models or retraining for different legal domains
Transparency	Clear traceability to source documents	Limited transparency; reasoning embedded in model weights
Customization	Highly customizable through knowledge base modifications	Limited to training data; requires retraining for modifications

The Legal Tech Context: impact on legal applications

The implications of choosing between RAG and fine-tuning are particularly significant in legal technology, where accuracy and reliability cannot be compromised. Legal professionals routinely engage with AI for document analysis, research, and content generation – tasks that demand both broad language understanding and deep domain expertise.

In the legal domain, RAG has emerged as a particularly compelling solution for several reasons. Law firms and legal departments can maintain their existing document management systems while leveraging them as knowledge bases for their AI applications. This approach ensures that AI-generated content remains grounded in verified legal sources and precedents.

The fine-tuning approach, while powerful for specific applications, presents unique challenges in the legal context. The dynamic nature of law, with constantly evolving precedents and regulations, means that fine-tuned models risk becoming outdated unless regularly retrained – a process that can be both expensive and time-consuming.

The Innovation frontier

Many successful legal tech implementations are now adopting hybrid approaches. For instance, using RAG for general legal research and document analysis, while employing fine-tuned models for specific, well-defined tasks like contract review or due diligence. This combination allows organizations to leverage the strengths of both approaches while mitigating their respective weaknesses.

Recent innovations in retrieval techniques have opened new possibilities for legal AI applications. Contextual Retrieval, a cutting-edge approach that preserves important document context during the retrieval process, has shown particular promise in legal applications where understanding the broader context of a legal provision or precedent is crucial.

Hybrid search algorithms that combine semantic understanding with traditional keyword-based approaches have also emerged as powerful tools for legal research and analysis. These systems can more effectively identify relevant legal documents and precedents, leading to more accurate and reliable AI-generated outputs.

Practical implications

For legal organizations implementing AI solutions, the choice between RAG and fine-tuning often comes down to practical considerations of scalability, maintenance, and accuracy. RAG systems offer the advantage of maintaining up-to-date knowledge through easily updated knowledge bases, while providing clear traceability back to source documents – a crucial requirement in legal applications.

Fine-tuning, while potentially more efficient for highly specialized tasks, requires careful consideration of the resources required for ongoing model maintenance and updates. The approach may be more suitable for specific, well-defined legal tasks where the underlying legal framework remains relatively stable.

Looking Ahead: tips for companies looking to use LLMs for legal tasks

The future of legal AI implementations will likely involve sophisticated combinations of both approaches, but organizations implementing RAG solutions should focus on three critical areas for success.

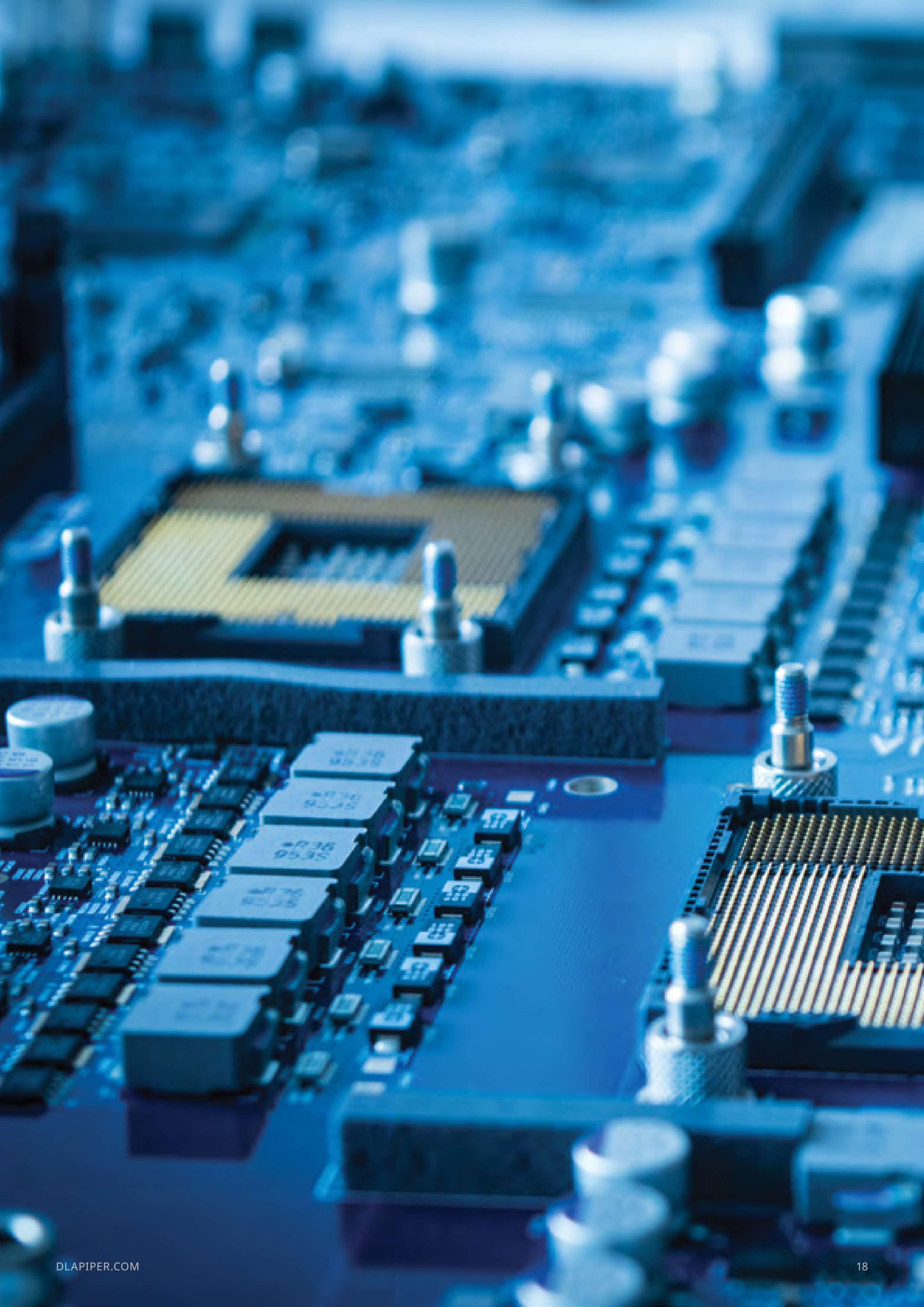
1) First, **careful architectural design and model selection are paramount**. Legal organizations must thoughtfully evaluate which models best suit specific tasks within their workflow. Some applications, such as basic document classification or metadata extraction, might benefit from lighter, more cost-effective models (even small language models). In contrast, complex tasks like contract analysis or legal research might require more sophisticated models capable of understanding nuanced legal contexts and choosing appropriate data sources (like the new series of complex reasoning models designed to spend more time "thinking" before responding).

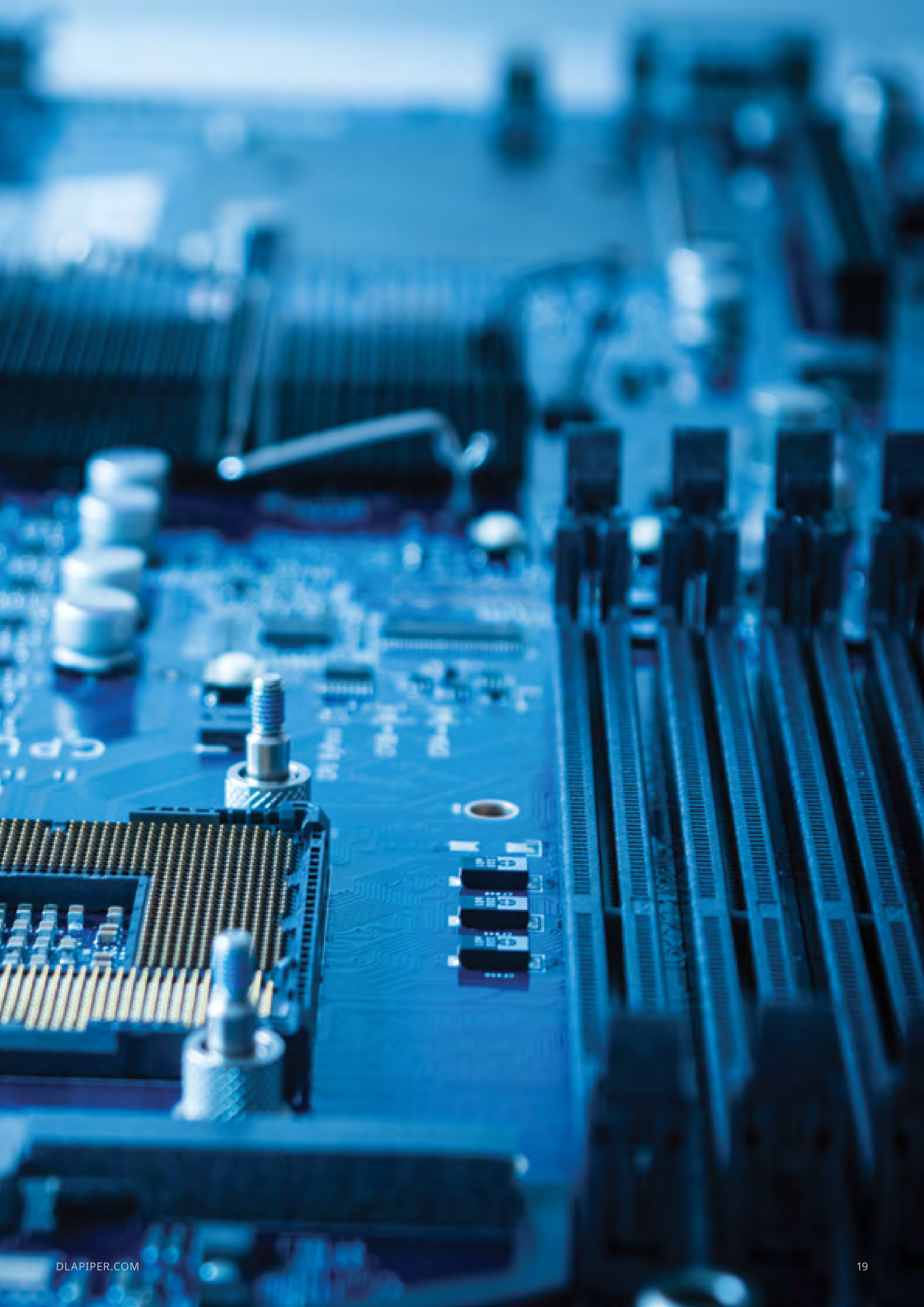
2) Second, the **focus must shift toward better data sources** rather than just raw computational power. The quality of training data and knowledge bases has emerged as a crucial differentiator in legal AI applications. High-quality legal content, well-documented precedents, and carefully curated internal knowledge bases are becoming as valuable as the AI models themselves. This emphasis on data quality follows the fundamental principle of "garbage-in, garbage-out" that becomes even more critical in legal applications where accuracy is paramount.

3) Third, **organizations should explore alternative methods to improve semantic search** and reduce costs. Even with increasingly powerful models, submitting entire legal documents to LLMs remains impractical and expensive. Advanced techniques like hybrid search algorithms, contextual retrieval, and sophisticated reranking systems are becoming essential tools in the legal AI toolkit. These approaches can drastically improve the relevance of retrieved information while maintaining cost-effectiveness and processing efficiency.

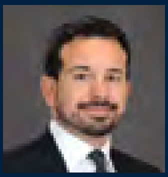
For organizations navigating the legal AI landscape, success hinges on carefully balancing technical innovation with practical implementation.
Rather than a one-size-fits-all approach, firms must weigh their specific needs – whether it's a boutique practice requiring rapid precedent retrieval or a corporate department needing broad multi-jurisdictional coverage – against their technical capabilities and resource constraints. This evaluation should guide the strategic implementation of RAG and fine-tuning technologies while ensuring alignment with both regulatory compliance and professional ethics.

The goal isn't merely to adopt cutting-edge technology, but to thoughtfully integrate AI systems that enhance legal practice while preserving the profession's core values of accuracy, confidentiality, and ethical conduct.



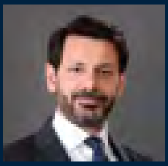


Contacts



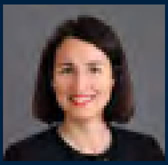
Giulio Coraggio
Partner
Head of Intellectual Property and Technology, Italy
T +39 02 80 618 1
giulio.coraggio@dlapiper.com

Gualtiero Dragotti
Partner
Global Co-Chair, Patent Group
T +39 02 80 618 1
gualtiero.dragotti@dlapiper.com



Alessandro Ferrari
Partner
Head of Technology Sector, Italy
T +39 02 80 618 1
alessandro.ferrari@dlapiper.com

Roberto Valenti
Partner
Head of Life Sciences Sector, Italy
T +39 335 73 66 184
roberto.valenti@dlapiper.com



Elena Varese
Partner
Co-Head of Consumer Good, Food and Retail Sector, Italy
T +39 02 80 618 1
elena.varese@dlapiper.com

Ginevra Righini
Partner
T +39 02 80 61 863 4
ginevra.righini@dlapiper.com

Marco de Morpurgo
Partner
Global Co-Chair, Life Sciences
T +39 06 68 880 1
marco.demorpurgo@dlapiper.com

Alessandro Boso Caretta
Partner
T +39 06 68 880 1
alessandro.bosocaretta@dlapiper.com

dlapiper.com

DLA Piper is a global law firm operating through various separate and distinct legal entities. Further details of these entities can be found at dlapiper.com.
This publication is intended as a general overview and discussion of the subjects dealt with, and does not create a lawyer-client relationship. It is not intended to be, and should not be used as, a substitute for taking legal advice in any specific situation. DLA Piper will accept no responsibility for any actions taken or not taken on the basis of this publication. This may qualify as "Lawyer Advertising" requiring notice in some jurisdictions. Prior results do not guarantee a similar outcome. Dec 13 2024 | A26259-3