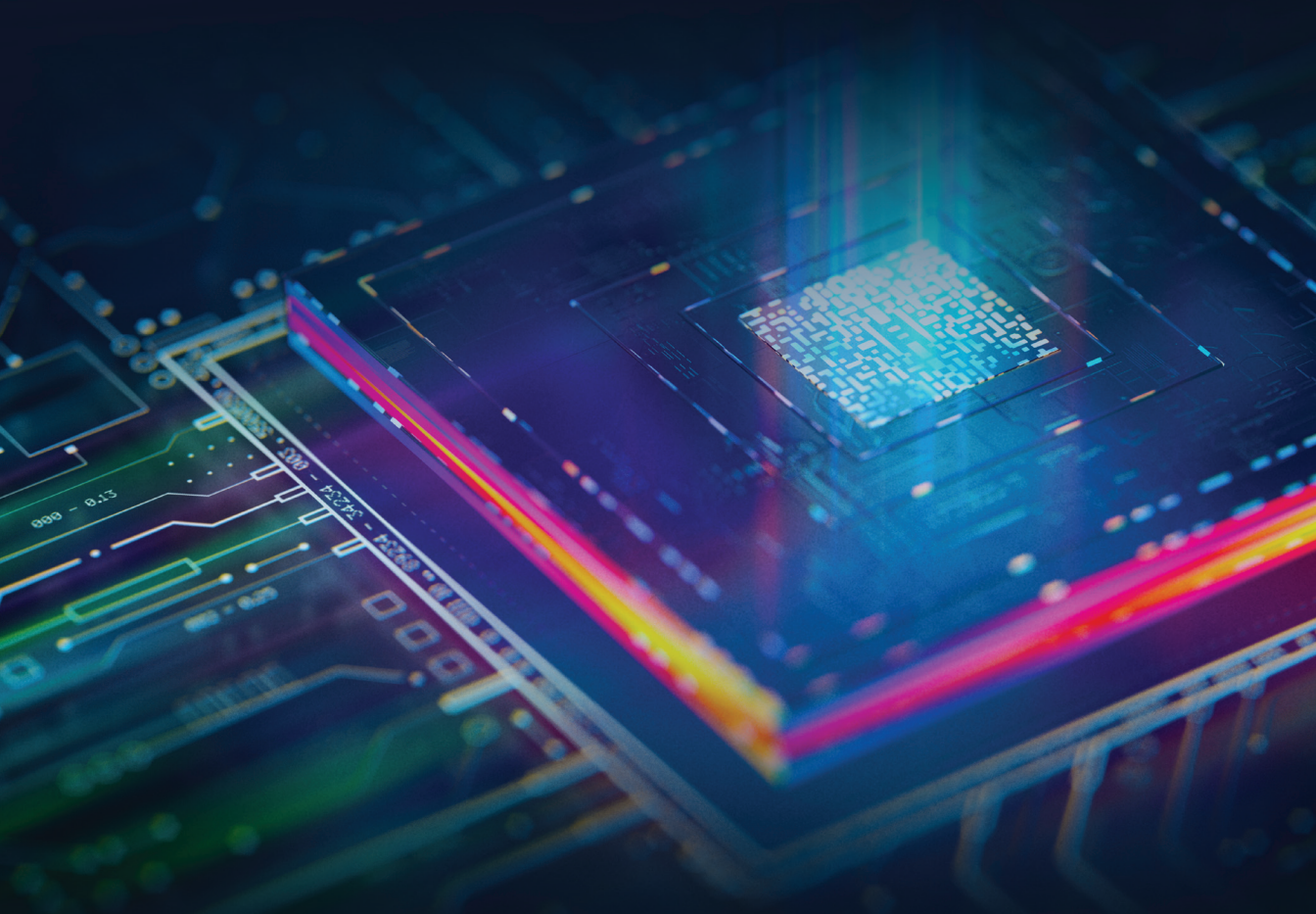
Providers of AI Systems

Ten Steps Towards Future Compliance

August 2023



Proskauer

Contents

Providers of AI Systems Ten Steps Towards Future Compliance

The global AI market was valued at $95.60 billion in 2021 and is predicted to reach $1.85 trillion by 2030, registering a compound annual growth rate of 32.9 per cent. Alongside this growth and the proliferation of AI use cases across industries, governments have quickly focused their sights on providers of the technology. Laws that specifically target AI system providers in the EU, UK and the US currently appear only in draft form or as non-binding guidance, or have otherwise not yet come into effect. However, the evolution of such draft laws and guidance into binding legislative instruments – and the coming into effect of the pending regulations – will soon alter this landscape forever.

AI system providers in the EU, UK and the US should start considering and implementing the expected basic requirements of future laws now. Such action will limit the scope of any re-engineering needed to achieve compliance with new laws once they come into force. It will also allow AI system providers to avoid cutting legal corners, and taking on unnecessary risk, in a rush to achieve compliance by applicable deadlines.

To assist in such exercise, this Proskauer briefing sets out ten steps that AI system providers can take towards compliance with the expected basic requirements of future laws. These steps are intended to be useful preliminary actions, rather than exhaustive steps to absolute compliance. The steps are followed by snapshots of certain key laws and guidance items, including their respective scopes, and known or anticipated dates of effect.

Step 1/
Cataloguing and Assessment

Catalogue your AI systems and conduct detailed assessments of each AI system throughout its life cycle to identify its potential risks and actual impacts (including any individual, social, economic, environmental and ethical impacts).

In particular:

a) during the initial design and development of your AI system, conduct impact assessments to identify and assess the potential impacts of the AI system;

b) following the deployment of your AI system, conduct impact evaluations to assess the actual impacts of the AI system and identify any unintended consequences; and

c) throughout the life cycle of your AI system, conduct risk assessments to identify and assess the potential risks (including technical, operational and security risks) associated with your AI system.

Pay particular attention to categories of persons or groups (including marginalised persons and vulnerable groups) that may be affected by your AI system.

Be sure to include in your organisation's risk register any risk items that you identify.

Step 2/
Risk and Impact Management

Implement, on an ongoing basis, technical measures (addressing the AI system itself) and organisational measures (addressing, e.g., governance) to manage and mitigate the impacts and risks that you have identified in Step 1.

For example, you might mitigate the risk of AI system 'drift' from a technical perspective by retraining or fine-tuning your AI system on new data, using real-time online learning and/or implementing ensemble methods (e.g., bagging, boosting or stacking).

Similarly, you might mitigate explainability and transparency risks from an organisational perspective by developing improved documentation procedures and creating easily understandable customer policies.

Establish reporting lines and an accountability framework that identifies the responsibilities of management and other staff in relation to AI system compliance, including which persons or teams are responsible for Steps 1–10.

Step 3/
Record-keeping and Traceability

In respect of each of your AI systems, document its key characteristics on an ongoing basis.

Key characteristics include each system's purpose; design and development process (including any substantial modifications); design specification and architecture; data sourcing and management; training and fine-tuning methodologies; testing protocols; capabilities and decisions; risk and impact assessments and evaluations; risk and impact mitigating measures; quality controls; and faults, failures and malfunctions.

Establish a quality management system to record how your AI system complies with applicable laws and ensure your AI system is traceable (e.g., by using a version control system to track changes to training data, or a data lineage tool to track data flows through the system).

Automate your record-keeping process where possible.

Step 4/
Bias and Training Data Qualities

Execute strategies to ensure your AI system does not discriminate against individuals or create unfair commercial outcomes.

Ensure your AI systems are trained and fine-tuned on high-quality, relevant datasets that are checked for errors, are representative and consider diversity factors like age, gender and ethnicity.

Consider using technical processes to reveal traits in datasets that most heavily influence decisions/outputs, and to highlight and remove sources of bias. For example, pre-process datasets to maintain as much accuracy as possible while reducing/removing any relationship between outcomes and protected characteristics. Alternatively, rebalance imbalanced datasets by adding or removing data about under/overrepresented subsets of the population.

Step 5/
Quality, Robustness, Accuracy and Cybersecurity

Implement an ongoing management system to maintain the quality and performance of your AI system.

This management system should:

a) ensure the legal compliance of your AI system;

b) ensure, to the extent possible, that your AI system can withstand unexpected events without failing or producing incorrect results. Ensure that it is robust to faults and inconsistencies by using redundancy solutions, e.g., backups;

c) ensure that your AI system is accurate in accordance with the generally acknowledged state of the art. This may involve using accuracy metrics and including levels of accuracy in your instructions documentation. Implement processes to assess and adjust the accuracy of outputs, including hyperparameter running;

d) ensure that your AI system is protected from unauthorised access and disruption. Identify and safeguard against AI-specific security incidents, including leakage of data, model inversions, data poisoning and prompt injections. Consider subscribing to security advisories to receive alerts of vulnerabilities and ensure patching processes are in place where components are externally maintained;

e) introduce real-time monitoring techniques to flag and trigger the remediation of incidents relating to items (a) – (d); and

f) allow end users to report inaccurate, biased or otherwise problematic outputs.

Step 6/
Explainability and Transparency

Ensure your AI system is explainable and not a ‘black box.’

In particular, ensure your AI system’s outputs and the rationale behind them are understood, that the capabilities and purpose of your AI system are communicated to relevant stakeholders, and that the mechanics of your AI system are explainable in human terms.

With a view to creating trust, inform users when your AI system is being used, provide clear and transparent information to users about how their data is being used and how your AI system works (including its capabilities and limitations), and provide users with an easy-to-use complaints-handling process if they wish to object to your operations.

For evidentiary purposes, use ‘model cards’ to record:

a) an overview of your AI system, including its purpose and key features;

b) a description of the data that was used to train and fine-tune your AI system, including its source and format;

c) an overview of your AI system’s training and fine-tuning processes, including the evaluation metrics used and hyperparameters tuned;

d) a summary of your AI system’s performance and accuracy; and

e) a description of any known or potential biases or limitations of your AI system.

Step 7/
Data Minimisation and Data Governance

Use detailed data governance processes to closely manage your data acquisition, collection, analysis, labelling, use, storage, aggregation and retention (as applicable).

When training, fine-tuning and operating your AI system, limit the collection of personal information to what is relevant and necessary to achieve your specific purpose.

Maintain and comply with your data protection policies and procedures, periodically reviewing your processing activities to ensure alignment with applicable requirements on lawful bases; fairness; transparency; storage limitations; purpose limitations; data minimisation; accuracy; security; impact assessments, automated decision-making; and other obligations under applicable data protection laws (e.g., UK GDPR, EU GDPR and applicable US state privacy laws).

Step 8/
Human Oversight

Construct mechanisms for human oversight over your AI systems to monitor ongoing performance, enable termination of hazardous operations and minimise harmful outcomes.

This could involve staff reviewing decisions and other outputs of your AI system, or having staff intervene in your AI system's operation.

Periodically test whether a human reviewer identifies an intentionally inaccurate decision or output, and maintain a log of all decisions and outputs that were overridden by a human reviewer and the reasons why.

Step 9/
Supply Chain Management

Where your AI system depends on the products or services of third parties, build compliance-focused safeguards into relevant supply contracts.

These safeguards should include additional technical and operational requirements that such third parties must satisfy (e.g., minimum security requirements).

Construct these safeguards to enable recovery of losses that you suffer as a result of compliance failures caused by the third parties.

Enable downstream providers to comply with their respective compliance obligations by preparing technical documentation and instructions of use.

Step 10/
Horizon Scanning

Determine, based on current draft laws, what future requirements you may need to satisfy.

In particular, determine whether you might provide a "high risk" AI system under the EU's AI Act; whether you are a type of "intermediary service" under the EU's DSA; and/or whether any sector-specific requirements might apply to you.

Plot a route to compliance now, satisfying in the short-term all applicable core compliance requirements under draft laws that are unlikely to change (noting that Steps 1–10 are not intended to be exhaustive in this regard).

Monitor timelines and updates to relevant laws as they progress to implementation.

Be cognisant of the requirement to appoint EU/legal representatives under the AI Act and DSA in applicable circumstances; it might be possible to use your existing EU GDPR representative for these purposes.

Adjust and develop internal policies, procedures and systems proactively to achieve full compliance by implementation dates.

Snapshots of Future EU, UK and US Federal Laws

EU and UK Regulatory Outlook

EU's Artificial Intelligence Act ("AI Act")

In-scope providers: The AI Act governs providers that place on the market, or put into service, AI systems in the EU (irrespective of whether those providers are physically present within the EU). Also within scope of the AI Act are providers of AI systems that are physically present outside the EU where the output produced by the AI system is used in the EU.

Research, testing, and development activities involving AI systems are exempt from the AI Act, as long as they respect fundamental rights and other applicable laws, and are not tested in real-world conditions. AI components provided under free and open-source licences are excluded from the AI Act, with the exception of foundation models.

Snapshot for providers: The AI Act adopts a risk-based approach to regulation, grouping AI systems into four risk categories: (1) AI systems that pose minimal/no risk (e.g., spam filters); (2) AI systems that pose a limited risk (e.g., chatbots and deep fakes); (3) AI systems that that pose a high-risk (e.g., systems intended to be used to make or materially influence recruitment decisions); and (4) AI systems that that pose an unacceptable risk (e.g., real-time biometric identification in publicly-accessible spaces). For each risk category, the AI Act specifies requirements for auditing, transparency and other obligations.

The bulk of the obligations under the AI Act fall on providers of high-risk systems. These include obligations to establish risk management systems; implement data governance practices; draw up appropriate technical documentation; maintain and facilitate record-keeping; provide transparency to users; implement processes that allow for human oversight; and ensure accuracy, robustness and cybersecurity. A provider's compliance with its obligations must be confirmed through a conformity assessment (sometimes involving an authorised body), with AI systems passing assessment required to bear the 'CE mark' before being placed on, or put into service in, the EU market. AI systems must also be registered in a public database and conformity assessments must be repeated in respect of AI systems that are the subject of substantial modifications (e.g., training data changes that significantly affect performance).

Sanctions under the AI Act can be up to €40 million or 6 per cent. of global turnover, whichever is higher.

Date of effect: The AI Act is currently in draft form and subject to change. European lawmakers hope to adopt it **before the end of 2023**, ahead of the European Parliament elections in 2024. The AI Act's core obligations are then expected to take effect after a **24-month grace period**.

EU's AI Liability Directive ("ALD")

In-scope providers: Providers of AI systems governed by the AI Act are also governed by the ALD.

Snapshot for providers: Due to the unique properties of AI systems (e.g., black box decision-making) there have been long-standing legal difficulties in proving causal links between the harmful outputs of AI systems and faults of AI system providers. Accordingly, the ALD updates national civil liability rules across the EU to make it easier for victims of AI-caused damage to prove who is liable and to receive compensation.

The ALD does this by requiring national courts to presume that the provider of an AI system is liable for damage where: (1) there is evidence of non-compliance with an EU or member state law intended to protect against the damage that occurred (e.g., the AI Act); (2) it can be considered reasonably likely that this non-compliance has influenced the relevant AI system's output or lack of output; and (3) the claimant has demonstrated that the AI system's output or lack of output caused the damage.

The ALD also empowers national courts to order the provider of a high-risk AI system to disclose information about its AI system, as part of a claim against the provider for damage suspected to have been caused by the AI system. Courts can

only order the disclosure of information which is necessary and proportionate to support the relevant claim. They must also consider whether trade secrets and confidential information of the AI system provider will be disclosed, and take measures to protect such trade secrets and confidential information.

Date of effect: The ALD is currently in draft form and subject to change. It is likely to become law **within the next 18 months**. Once negotiated and adopted, EU Member States will be required to transpose the terms of the ALD into national law, **likely within 2 years**.

EU's Digital Services Act ("DSA")

In-scope providers: The DSA governs providers that offer online intermediary services in the EU (irrespective of whether those providers are physically present within the EU). As such, it applies to all digital services that connect consumers to goods, services, or content. Intermediary services are those that offer network infrastructure such as internet access providers, domain name registrars and include hosting services that in turn includes online platforms/marketplaces. This might include certain AI providers. There are additional obligations on very large online platforms (VLOPs) and very large online search engines (VLOSEs) that reach at least 45 million active monthly users in the EU. The European Commission has designated certain platforms/engines as VLOPS/VLOSEs.

Snapshot for providers: The DSA's aim is to create a safer digital space. Its requirements depend on what the provider is, with cumulatively more burdensome requirements considering the role, size and impact of the provider's service/platform. The requirements on all providers include: providing transparent information in terms and conditions in relation to content moderation; having a designated point of contact/s or legal representative/s in the EU to communicate with relevant authorities/users; and publishing a content moderation report. For all "hosting services" (including all "online platforms"), additional obligations include having processes in place for users to report illegal content. There are then further requirements for all "online platforms" (unless the small enterprise exemption applies), including: having a complaint and redress mechanism in place as well as requirements in relation to transparency of advertising on online platforms. For VLOPs and VLOSEs there are additional obligations.

Sanctions under the DSA could potentially be up to 6 per cent. of global annual turnover.

Date of effect: The rules apply to "intermediary services," "hosting services" and "online platforms" from **17 February 2024**. For VLOPs and VLOSEs, the rules apply from **25 August 2023**.

EU's Digital Markets Act ("DMA")

In-scope providers: The DMA affects "gatekeeper" platforms (whether or not based in the EU). Alphabet, Amazon, Apple, ByteDance, Meta, Microsoft and Samsung have self-declared as meeting the thresholds in the DMA. The EU Commission is soon to confirm whether these businesses should be considered gatekeeper platforms and whether any other business should be categorised as such.

Snapshot for providers: The DMA sets out a list of dos and don'ts for gatekeepers, with the aim of ensuring "fair and open digital markets" and levelling the playing field for other third-party providers. This means certain providers who are not gatekeepers might benefit in their use of gatekeeper products/services and should watch out for these changes/benefits.

Gatekeepers must not, for example: (1) require app developers to use gatekeeper services to appear in app stores; (2) rank gatekeeper products/services in a preferential way to other third parties; or (3) use data of business users when gatekeepers compete with them on their platform. On the other hand, gatekeepers must, for example: (1) allow business users to promote their services and conclude contracts outside of the gatekeeper platform; (2) allow business users to access data generated by their use of the gatekeeper platform; and (3) allow third parties to inter-operate with the gatekeeper's own services, including making certain functionalities available to third parties so as to enable interoperability of messenger services.

Sanctions under the DMA are up to 10 per cent. of global annual turnover or up to 20 per cent. of global annual turnover for repeated infringements.

Date of effect: Once the EU Commission has confirmed the identify of gatekeeper platforms (**expected by early September 2023**), those gatekeepers will have **6 months to comply** with the DMA rules.

UK's AI-focused Publications

In-scope providers: The UK's AI-focused publications target businesses developing and using AI in the UK, as well as businesses outside the UK that provide products or services using AI to customers in the UK.

Snapshot for providers: The UK government intends to create a world-leading ecosystem for AI development and deployment, as outlined in, among other things, the UK's National AI Strategy, Roadmap to an Effective AI Assurance Ecosystem, and AI Action Plan. Rather than propose any laws specifically targeted at AI providers, various individual departments and regulatory bodies have published papers and strategies. Examples include: (1) the ICO's guidance paper on the AI Auditing framework, explaining decisions made with AI, and AI and data protection; (2) the DCMS, DBEIS and Office for Artificial Intelligence's policy paper on establishing a pro-innovation approach to regulating AI; (3) the DSIT and Office for Artificial Intelligence's follow-up white paper on a pro-innovation approach to regulating AI; and (4) the CDEI's portfolio of AI assurance techniques.

The papers and strategies outline various guiding principles, checklists and techniques aimed at ensuring the responsible provision of AI systems. For example, providers are encouraged to implement AI systems that are consistent with principles of safety, security, robustness, transparency, explainability, fairness, accountability, governance, contestability and redress. They also emphasise the need for appropriate documentation and record-keeping, audit and evaluation processes (including impact assessments and evaluations, and risk assessments) and decision-mapping exercises. They further outline the future importance of trust and assurance mechanisms, including conformity assessments and certifications.

The UK government's expectation is that individual regulators (e.g., the ICO, the FCA, the CMA's Digital Markets Unit) will use the proposed principles, checklists and techniques as a basis to provide further guidance and, where relevant, sanctions and enforcement within their respective remits.

Date of effect: **The various papers and strategies are already in place.** While they do not contain binding controls, they are **likely to influence how regulators interpret and apply existing legal and regulatory requirements** that govern AI providers. UK political parties have also signalled an intention to build upon the papers and strategies by introducing formal regulation in the future (such as in the context of regulating AI ethics and liability).

US Federal Regulatory Outlook

The US lags considerably behind the EU and UK when it comes to AI regulation. The 118th US Congress, the current meeting of Federal legislators, ends on 3 January 2025, and it is not clear whether any AI laws will be passed before then. However, the following proposals are worth note:

Senate Majority Leader Charles Schumer's AI Framework

In-scope providers: Unclear.

Snapshot for providers: In April, Senate Majority Leader Schumer announced his intention to spearhead efforts to craft a regulatory framework for AI, centred around requiring disclosures and transparency in connection with AI systems' development, training data sources, and technical functionality. However, no timeline or legislative details have been provided.

Date of Effect: Unclear.

Senate Democrats' Algorithmic Accountability Act

In-scope providers: AI providers within the jurisdiction of the Federal Trade Commission (FTC) (i.e., AI providers who operate in the US).

Snapshot for providers: This bill was introduced by Democratic legislators during the 117th US Congress. Although the bill expired with the end of that Congress, it may be reintroduced in this current Congress. The bill would require AI providers to perform impact assessments (detailing their testing and evaluation efforts, effects on consumers, employee training, user guardrails, etc.) and provide reports to the FTC both before and after deployment. Although this bill could be reintroduced in the current Congress, it did not find significant bipartisan support in the 117th Congress, and Democrats may opt to roll its concepts into Senator Schumer's broader effort.

Date of Effect: Unclear.

Senate Democrats' Algorithmic Justice and Online Platform Transparency Act

In-scope providers: AI providers who offer "online platforms" and are within the jurisdiction of the FTC.

Snapshot for providers: This bill would prohibit AI providers from offering tools that discriminate on the basis of protected characteristics (race, gender, etc.) and would require disclosures and reporting with respect to online platforms' use of "algorithmic processes," data collection practices, and content moderation practices. Like the Algorithmic Accountability Act, it would be enforced by the FTC. Also like that bill, this bill was introduced by Democratic legislators during the 117th US Congress but did not find significant bipartisan support, and Democrats may opt to roll its concepts into Senator Schumer's broader effort.

Date of Effect: Unclear.

Other Executive and Regulatory Agency Efforts

Although Federal AI legislation does not appear imminent, other elements of the US government are making efforts to protect US consumers without waiting for explicit Congressional AI-specific directives. For example, the FTC is currently investigating whether OpenAI "engaged in unfair or deceptive privacy or data security practices or engaged in unfair or deceptive practices relating to risks of harm to consumers," and several major AI providers recently met with President Joseph Biden and agreed to voluntary "security testing, in part by independent experts; research on bias and privacy concerns; information sharing about risks with governments and other organizations; development of tools to fight societal challenges like climate change; and transparency measures to identify AI-generated material." Although these efforts are ad hoc, they serve as a reminder that AI providers remain subject to non-AI-specific rules and regulations and demonstrate that US Federal executive and independent authorities do not necessarily need to wait for legislation before taking action to protect consumers.[1]

[1] A handful of state and local jurisdictions within the US have enacted, or are considering enacting, laws pertaining to specific uses of automated tools, including AI. To date, these laws fall into one or both of the following categories: (a) laws governing use of automated tools in an employment context, and (b) laws prohibiting use of automated tools to make decisions based on protected personal traits (e.g., race, gender). The most well-known of these laws, which falls into both of the aforesaid categories, is New York City's Local Law 144, which requires employers and employment agencies who use automated tools in the hiring context to subject those tools to annual "bias audits", and also imposes notice and opt-out requirements. Massachusetts is the only state whose legislature is currently considering a specific law to regulate AI providers generally (as opposed to AI users) (SB31), but this bill is in very early stages and thus may change significantly before passage or may not pass at all. Note also that Federal legislation, if enacted, may preempt some or all of any then-existing state or local AI laws.

Key Contacts

Proskauer's lawyers are experts in AI law, policy and practice.

We regularly advise on the development and deployment of AI technologies, enabling businesses on both sides of the AI market to manage risk issues and execute key strategies. The firm's global team is recommended in the professional directories for its "niche in robotics and artificial intelligence."

For further information on the matters highlighted in this briefing or for assistance with any AI project, please contact one of the following team members or your usual Proskauer contact.

Oliver R. Howley
Partner
ohowley@proskauer.com

Kelly M. McMullon
Special Counsel
kmcmullon@proskauer.com

Peter J. Cramer
Associate
pcramer@proskauer.com

This material is for general information only and is not intended to provide legal advice.

Proskauer