



Crisis Management in the Age of AI: Smarter Red Teams and AI Tools for Anticipating Corporate Catastrophes

Corporate, Securities, and Governance

Technology, Privacy, and eCommerce



Banner artwork by [showcake / Shutterstock.com](#)

The COVID-19 pandemic taught all of us about crisis management in real time. Working from our home offices and our dining room tables, we authored a series of articles about strategies for not just responding to business crises, but getting ahead and anticipating them. Now, as AI reshapes the legal industry in ways that feel just as seismic, we are returning to this topic with a new three-part series on how these technologies can augment and transform the practice of corporate crisis management.

The risk landscape has only grown more complex. Crises move faster, spread farther, and carry steeper reputational and legal consequences than ever. Some of those risks stem directly from AI itself. But AI has also matured from an abstract concept into a practical toolkit that in-house counsel and crisis teams can put to work right now. This series sits at the intersection of those two realities, offering in-house legal professionals concrete, actionable guidance on weaving AI into their crisis preparedness and response programs.

This first installment zeroes in on the anticipation phase, focusing on how AI tools can sharpen red teaming, horizon scanning, and premortem exercises so organizations can spot and prepare for threats before they materialize. In our 2020 series, we explored red teaming and premortems as foundational tactics: structured methods for challenging assumptions, exposing blind spots, and stress-testing corporate plans for hidden vulnerabilities. We also looked at how crisis teams should be assembled and prepared before disaster strikes, and how multinational crises demand sensitivity to culture, language, and local legal frameworks. Those strategies are still essential. What has

changed is that AI tools can now amplify their reach and address their inherent limitations.

For in-house counsel and their business partners, the question is no longer whether AI has a role in crisis anticipation. It is how to deploy it responsibly, harnessing its analytical power while preserving human judgment, protecting privilege, and maintaining accountability.

The limits of human red teaming and where AI may expand the aperture

Red teaming, at its core, is an exercise in structured contrarian thinking. The technique traces its modern corporate lineage to intelligence and military applications, most notably the CIA's Red Cell, created after the September 11 attacks to challenge prevailing analytical assumptions. Adapted to the corporate setting, red teaming enables organizations to identify potential crises, stress-test existing policies, and develop contingency plans by deliberately questioning what others take for granted.

Red teaming enables organizations to identify potential crises, stress-test existing policies, and develop contingency plans by deliberately questioning what others take for granted.

The concept has also gained traction as an AI governance mechanism in its own right. Governments around the world have begun converging on red teaming as a primary strategy for managing the risks of generative AI, applying structured adversarial testing to identify flaws and vulnerabilities in AI systems before deployment. (Andrew Burt, "How to Red Team a Gen AI Model," Harvard Business Review (Jan. 2024)). Just as red teaming has become a best practice for evaluating AI systems themselves, AI tools can in turn strengthen the red teaming exercises that organizations use to anticipate corporate crises.

The method's effectiveness depends on the capability, independence, and creativity of the team assembled. Yet even well-constructed human red teams face structural constraints that limit their output:

- **Normalcy bias.** Research suggests 70 percent of people experience normalcy bias, or the cognitive tendency to expect the future to resemble the past, during crisis situations. (John Leach, "Why people 'freeze' in an emergency: temporal and cognitive constraints on survival responses," *Aviation, Space, and Environmental Medicine* (2004).) Red team participants are not immune. Despite instructions to think adversarially, team members may unconsciously anchor to historical precedents and discount genuinely novel threat vectors.
- **Groupthink.** Even in teams designed to foster dissent, social dynamics can suppress unconventional ideas. Junior members may defer to senior participants. Consensus can emerge prematurely, narrowing the range of scenarios explored.
- **Knowledge gaps and time constraints.** Human teams are limited by their collective expertise and the hours available for the exercise. A red team composed of lawyers, communications professionals, and operations executives will naturally generate scenarios within those domains, potentially overlooking threats from unfamiliar technical, geopolitical, or regulatory areas.

- **Insufficient scenario diversity.** Traditional exercises tend to produce a manageable but inherently limited set of scenarios. When a team generates ten or fifteen threat narratives, those narratives inevitably reflect the participants' own mental models and professional experiences rather than the full universe of plausible risks.



AI will not automatically eliminate these constraints when introduced into red teaming exercises. But it can meaningfully expand the aperture of analysis. Large language models and scenario-generation tools can produce hundreds of threat narratives in the time a human team would develop a dozen. These technologies can draw on training data spanning industries, geographies, and historical periods unfamiliar to any individual participant. And because AI outputs are not subject to social pressure, they can surface uncomfortable or counterintuitive scenarios that human participants might self-censor.

The appropriate model is not, however, replacement. In this context, like many others, AI is a tool of augmentation in this context and in many others. AI-generated scenarios serve as inputs to human deliberation, or raw material that the red team evaluates, refines, and prioritizes using contextual judgment that AI cannot replicate. The human team in an AI-enabled red team exercise retains responsibility for assessing plausibility, organizational relevance, and strategic significance. AI broadens the initial funnel, but human experience and expertise shape the output into actionable intelligence.

AI broadens the initial funnel, but human experience and expertise shape the output into actionable intelligence.

In practice, the division of labor might look like this: Before a red team convenes, an AI tool is prompted to generate a broad set of threat scenarios tailored to the organization's industry,

geographic footprint, and current strategic initiatives. The red team reviews these outputs alongside its own independently generated scenarios, identifies themes and gaps, and selects a subset for deeper exploration. The AI's contribution ensures that the team's starting universe of risks is broader than what any group of individuals could independently produce, but the most meaningful conclusions will still be tied to human effort, judgment, and innovation.

AI-powered horizon scanning for continuous vigilance

If red teaming is the periodic exercise of adversarial imagination, horizon scanning is its continuous counterpart. Horizontal scanning involves ongoing surveillance of an organization's threat environment for signals that a crisis may be developing. The most sophisticated crisis management planning involves not just preparing for an immediate crisis but anticipating crises before they happen. AI's capacity to process vast volumes of unstructured data in real time makes it particularly well-suited to this function.

The most sophisticated crisis management planning involves not just preparing for an immediate crisis but anticipating crises before they happen.

AI-powered monitoring platforms can simultaneously scan local, national, and international news feeds, social media activity, regulatory filings, litigation databases, patent registrations, earnings call notes, government procurement records, and other open-source signals. Natural language processing enables these tools to identify patterns. They can detect subtle shifts in sentiment, track the emergence of new regulatory language, identify supply-chain vulnerability indicators, and flag correlations that might not register with a human analyst reviewing any single data stream in isolation.

[Watch the Latest Legal Innovation Spotlight with ACC President & CEO Jason L. Brown and Ted Theodoropoulos, CEO of Infodash, on AI and the Future of Legal Services.](#)

For example, NLP tools can track how regulators describe an industry's practices over successive enforcement actions or guidance documents. A gradual shift from descriptive to prescriptive language, or an uptick in references to specific statutory provisions, could signal forthcoming regulatory action. Similarly, monitoring litigation filings across jurisdictions can reveal emerging theories of liability before they coalesce into coordinated litigation campaigns.

The value of AI-assisted early warning becomes clearer against the shortcomings of conventional monitoring. For example, a global manufacturer with complex multi-tier supply chains might not

detect that a critical sole-source supplier is facing financial distress until a missed shipment disrupts production. An AI monitoring tool scanning the supplier's public filings, credit signals, local news coverage, and employee review platforms could flag deteriorating conditions weeks or months earlier, providing time to qualify alternative suppliers or build buffer inventory.

Similarly, a company facing potential product-safety claims might not recognize the scope of consumer complaints until a regulatory investigation is announced. AI tools monitoring consumer forums, social media posts, and adverse-event databases, on the other hand, could identify clustering patterns and escalating complaint volumes before formal regulatory attention.

Horizon scanning should not operate in isolation from the other anticipatory disciplines discussed in this article. Signals identified through AI-powered monitoring should feed directly into red teaming exercises and premortem analyses, creating a continuous feedback loop in which emerging threats detected in real time inform the scenarios that teams explore in structured exercises. This integration ensures that an organization's crisis-anticipation framework functions as a cohesive system rather than a collection of disconnected activities.

These are not hypothetical capabilities. They represent the operational logic of commercially available AI technologies already in use across industries. The question for legal and compliance teams is not whether such tools exist, but how to integrate them into existing enterprise risk management programs.

Achieving practical integration of AI into red team exercises

Effective deployment of these technologies in the red team and larger crisis management contexts requires more than purchasing a monitoring platform. Organizations should address several operational questions, including:

- **Vendor selection.** Evaluate how the platform handles data storage, retention, and security. Distinguish between open-loop systems (which may use organizational inputs to train the model) and closed-loop systems (which do not). Courts and regulators increasingly hold the deploying organization, not the model provider, accountable when AI tools cause harm, making vendor due diligence a matter of direct legal exposure. (M. Alejandra Parra-Orlandoni & Paulo Carvão, "You Outsourced the AI — but You Still Own the Risk," Harvard Business Review (July 2026)). Assess whether the platform's NLP capabilities are trained on data relevant to your industry and regulatory environment.
- **Integration with enterprise risk management.** AI monitoring outputs should feed into existing ERM governance structures rather than operating as a standalone function. Define escalation pathways: which alerts require immediate legal review, which should route to the crisis-management team, and which are informational only.
- **Alert triage.** Unfiltered AI monitoring will generate significant noise. Invest in calibrating alert thresholds, defining materiality criteria, and establishing triage protocols that prevent alert fatigue while ensuring genuine signals receive prompt attention.
- **Human oversight.** AI monitoring is a supplement to, not a substitute for, human analytical judgment. Alerts should be reviewed by personnel with sufficient context to evaluate their significance. Automated systems should flag; humans should decide.

Reimagining the premortem with Generative AI

The premortem exercise, or imagining that a plan has already failed spectacularly and working backward to identify what went wrong, is among the most powerful tools available for surfacing hidden assumptions and testing organizational plans for vulnerabilities. Psychologist Gary Klein developed the technique to combat the overconfidence that typically accompanies plan adoption: by instructing participants to assume failure has already occurred, the exercise breaks emotional ownership of the plan and enables genuinely critical analysis. (Gary Klein, *Performing a Project Premortem* (Harvard Business Review, 2007).)

In traditional premortems, participants individually generate reasons the plan might have failed, then consolidate and discuss those reasons as a group. The exercise is deliberately rapid, often completed in under an hour, and produces a prioritized list of vulnerabilities that informs plan revision.

Generative AI adds a new dimension to this practice. Where a human team might generate fifteen to twenty failure reasons in a session, a large language model can produce detailed, multi-path failure narratives that explore cascading consequences, second-order effects, and interactions between failure modes that participants might not independently identify.

AI as a failure-narrative generating partner

An AI tool prompted with the details of a corporate plan, such as a product launch, a restructuring, an acquisition integration, can generate dozens of distinct failure scenarios, each with a plausible causal chain. These narratives can explore failures originating from different domains: regulatory, operational, reputational, competitive, technological, and geopolitical. They can model how failures in one domain cascade into others, producing the kind of multi-system breakdowns that characterize genuine corporate catastrophes.

The output serves not as a finished analysis but as rich input for board-level risk discussions, management planning sessions, and legal risk assessments. Decision-makers can evaluate which AI-generated failure modes they find credible, which they have already accounted for, and which represent blind spots warranting further investigation.

The output serves not as a finished analysis but as rich input for board-level risk discussions, management planning sessions, and legal risk assessments.

This kind of structured failure analysis at the board level is increasingly urgent. Boards must consider at least six categories of AI disruption scenarios, ranging from extreme operational changes to existential threats that could render an organization's business model obsolete. (Vivek Sharma & David C. Edelman, "6 Ways AI Could Disrupt Your Business," Harvard Business Review (Jan. 2024)). AI-assisted premortems can help leadership teams systematically work through these categories rather than relying on ad hoc assessments.

Maintaining privilege and work-product protections

AI-assisted premortems and red teaming exercises raise important questions about privilege and work-product protection. These analyses may qualify for protection under the attorney-client privilege

(if conducted for the purpose of obtaining legal advice) or the work-product doctrine (if conducted in anticipation of litigation or in preparation for a potential legal proceeding). To strengthen potential privilege claims, organizations should structure AI-assisted premortems with the following considerations in mind:

1. Conduct the exercise under the direction of counsel, with counsel framing the prompts and guiding the analysis.
2. Document the legal purpose of the exercise — for example, assessing legal risks associated with a proposed course of action or preparing the organization to respond to potential litigation or regulatory inquiry.
3. Limit distribution of AI-generated outputs and final analyses to those with a need to know.
4. Ensure that the AI platform used does not retain, share, or train on the inputs provided. If the platform operates on an open-loop basis, privilege may be compromised by the disclosure of confidential information to a third-party system without adequate safeguards.
5. Maintain clear records distinguishing legal analysis from business planning, recognizing that dual-purpose documents may receive less robust protection depending on the jurisdiction.

Structuring an exercise to support privilege does not guarantee protection. Courts evaluate privilege claims based on the totality of the circumstances, and the law governing AI-generated materials in this context is still developing. The prudent approach is to structure exercises conservatively while recognizing that privilege outcomes will ultimately depend on judicial interpretation as this area of law matures. In-house counsel directing AI-assisted exercises should also be mindful of their professional ethical obligations, including duties of competence, supervision, and confidentiality, as addressed in evolving guidance from the American Bar Association and state bar associations regarding lawyer use of AI tools.

The Sedona Conference is actively developing authoritative guidance in this area. In the 2026 Commentary on Generative AI, the Attorney-Client Privilege, and the Work-Product Doctrine, which will address how privilege and work-product analyses apply to GenAI data generated in the corporate context, including by employees and in-house counsel using GenAI tools in connection with legal advice and anticipated litigation. The Sedona Conference has also published a Commentary on the Application of Attorney-Client Privilege and Work-Product Protection to Documents and Communications Generated in the Cybersecurity Context, which provides an analogous analytical framework for technology-assisted activities conducted under counsel direction. These resources will be essential references for organizations structuring AI-assisted crisis-anticipation exercises with privilege protection in mind.

Second Edition

Cybersecurity Toolkit

for In-house Lawyers

DOWNLOAD TODAY ↓

Governing AI in the anticipation phase

Integrating AI into crisis-anticipation workflows introduces its own category of risk. Governance failures in the anticipation phase can compound rather than reduce organizational vulnerability. AI has made the cybersecurity and risk environment increasingly “brittle, anxious, nonlinear, and incomprehensible,” rendering traditional planning frameworks insufficient. (Hise O. Gibson, “AI Is Reshaping Cyber Risk. Boards Need to Manage the Threat,” *Harvard Business Review* (Apr. 2026)). That observation applies with equal force to crisis anticipation: organizations that integrate AI tools without corresponding governance structures risk introducing new vulnerabilities rather than reducing existing ones.

Over-reliance and algorithmic normalcy bias. Much like human teams, AI systems can exhibit their own version of normalcy bias. Models trained predominantly on historical data may systematically underweight genuinely novel threats — the very category of risk that crisis anticipation most needs to address. If an organization relies too heavily on AI-generated threat assessments without testing them against human judgment and lived experience, it may develop a false sense of security anchored to the AI’s training distribution rather than the actual threat environment. The antidote is disciplined human oversight. AI outputs should be treated as one input among several — not as authoritative threat assessments. Red teams evaluating AI-generated scenarios should be explicitly tasked with identifying what the model might have missed, what assumptions underlie its outputs, and where its training data may create blind spots.

Data privacy and confidentiality. Using AI tools for crisis anticipation often requires inputting sensitive organizational information, such as strategic plans, financial projections, supply-chain details, personnel data, and pending litigation information. When this information is provided to third-party AI platforms, organizations face meaningful data-privacy and confidentiality risks. The critical threshold question is whether the platform operates on a closed-loop or open-loop basis. Closed-loop systems process inputs without retaining them or using them to train the underlying model. Open-loop systems may retain inputs, use them for model improvement, or store them in ways that create

risk of inadvertent disclosure. For sensitive organizational information, the distinction is material. Before deploying any AI tool in the crisis-anticipation function, organizations should evaluate: Where is data stored? Is it encrypted in transit and at rest? Does the platform comply with applicable data-residency requirements? Is data used to train the model or shared with third parties? What deletion mechanisms exist, and are they verifiable? These bear directly on whether the organization can maintain confidentiality obligations owed to clients, business partners, and employees.

Regulatory compliance. Organizations should also recognize that AI tools deployed for crisis anticipation may themselves fall within the scope of emerging AI-specific regulation. The European Union's AI Act, which began phased implementation in 2024, imposes transparency, risk-assessment, and human-oversight requirements on certain categories of AI systems. In the United States, states including Colorado and Illinois have enacted or proposed legislation governing automated decision-making and AI-driven risk assessments. At the federal level, evolving SEC disclosure expectations may require organizations to address AI-related risks in their public filings. In-house counsel should monitor these developments as part of the broader governance function, ensuring that AI tools used in crisis anticipation comply with applicable legal requirements as they take effect.

Human oversight and escalation protocols. Governance of AI in the anticipation phase requires clear protocols for human oversight and escalation. At minimum, organizations should establish:

- Defined roles specifying who is responsible for reviewing AI outputs, validating their relevance, and deciding on escalation.
- Escalation criteria establishing when an AI-generated alert or scenario must be elevated to senior leadership, the crisis-management team, or the board.
- Periodic review mechanisms to audit AI tool performance, recalibrate alert thresholds, and assess whether the tools remain fit for purpose as the threat environment evolves.
- Documentation protocols to maintain records of AI-generated outputs, human decisions made in response, and the reasoning underlying those decisions — creating an audit trail that demonstrates responsible governance.

Practical takeaways for in-house counsel

AI tools will not replace the human judgment, institutional knowledge, and strategic sense-making that effective crisis anticipation demands. They can, however, expand the range of threats organizations detect, accelerate the speed at which signals are identified, and deepen the analytical rigor of exercises designed to stress-test corporate plans. Achieving these benefits, however, requires cross-functional collaboration among legal, IT, information security, compliance, and executive leadership; in-house counsel's role is to establish the legal guardrails within which these teams operate.

For organizations seeking to integrate AI into their anticipatory frameworks, four guiding principles provide a starting point:

1. Deploy AI-generated scenario analysis as an input to — not a substitute for — human red team deliberation. Use AI to broaden the funnel of risks considered; rely on human expertise to prioritize and contextualize.

-
2. Integrate AI monitoring into existing enterprise risk management governance structures, with defined escalation pathways, triage protocols, and human review at each decision point.
 3. Structure AI-assisted premortem exercises under counsel direction to support potential privilege claims, while recognizing that privilege protection in this context is fact-dependent and legally still developing.
 4. Govern AI tools with the same discipline applied to any other component of the risk-management function: vet vendors rigorously, protect confidential information, maintain human oversight, and stay current with regulatory developments.

The integration of AI into crisis anticipation is not a theoretical proposition but a present operational opportunity. The task for legal and corporate leaders is to ensure that this integration is disciplined, accountable, and aligned with the governance standards their organizations already apply to high-stakes decision-making.

The next installment in this series will examine AI's role in real-time crisis response: how organizations can leverage and deploy AI tools during active emergencies to coordinate multidisciplinary teams, manage communications across jurisdictions, and make time-sensitive decisions under conditions of uncertainty, including the use of AI-assisted triage during active litigation and regulatory investigations. The final installment will address the post-crisis phase, exploring how AI can support after-action reviews, regulatory remediation, and organizational learning.

[Join ACC for more AI insights!](#)

Disclaimer: The information in any resource in this website should not be construed as legal advice or as a legal opinion on specific facts, and should not be considered representing the views of its authors, its authors' employers, its sponsors, and/or ACC. These resources are not intended as a definitive statement on the subject addressed. Rather, they are intended to serve as a tool providing practical guidance and references for the busy in-house practitioner and other readers.

[Sean Kennedy](#)



Senior Director, Head of US Ethics & Compliance

Moderna

Sean Kennedy is Senior Director and Head of U.S. Ethics, Compliance & Privacy at Moderna and a member of its U.S. Leadership Team. A healthcare attorney and compliance leader, he advises senior leaders on commercialization, healthcare compliance, privacy, and business risk. His work includes building practical compliance programs and developing AI-enabled tools and workflows to make guidance more accessible, improve review processes, and support sound business decisions. Before joining Moderna, Kennedy held several legal and compliance leadership roles at Bristol Myers Squibb, including Senior Director of Healthcare Compliance Strategy & Operational Risk Management and Senior Counsel in Worldwide Commercialization & Development Law. His

experience spans oncology, product development and launches, patient support programs, patient advocacy, corporate affairs, and enterprise compliance strategy. Across these roles, he has helped teams translate complex legal and regulatory requirements into practical approaches that support business goals and patient needs. Earlier in his career, he practiced litigation at Faegre Drinker and Carroll McNulty & Kull, including pharmaceutical multidistrict litigation and bellwether trials, and served as a judicial law clerk in New Jersey Superior Court. He earned his J.D. from Seton Hall University School of Law and his B.A. in Political Science, cum laude, from Seton Hall University. He is admitted to practice in New York and New Jersey.

[Kaitlyn Stone](#)



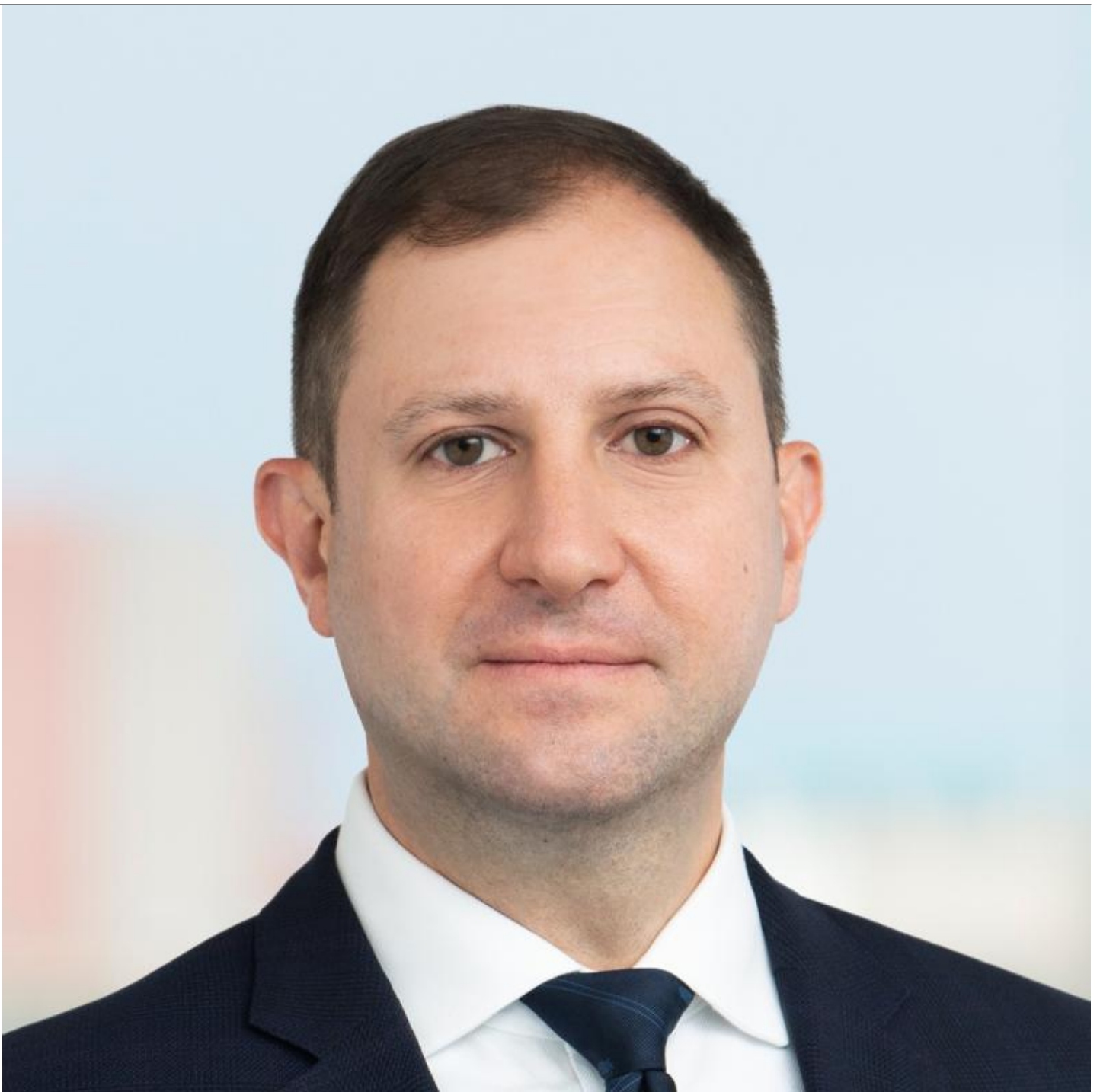
Litigation Partner & Co-Chair of Artificial Intelligence Practice

Barnes & Thornburg LLP

Kaitlyn E. Stone is a litigation partner and co-chair of the Artificial Intelligence practice group in the New Jersey office of Barnes & Thornburg. Stone defends companies in a variety of high-stakes litigation. She counsels leading health and life sciences, pharmaceutical, and medical device companies in product liability and mass tort matters, with an emphasis on prescription medications, medical devices, and consumer goods. Her areas of focus span multidistrict litigation (MDL), class actions, and coordinated state proceedings involving thousands of plaintiffs, to single-plaintiff disputes. Stone is also a formidable advocate in complex commercial and real estate litigation, with extensive experience in business tort, land use, environmental, and public contract cases. Stone lectures and publishes extensively on issues related to product liability, crisis management, and AI. She is an author of recent publications featured in Bloomberg Law, Law360, and more. She earned

her B.A. from the University of Rhode Island and her J.D. from Seton Hall University School of Law. She previously served as law clerk to the Honorable Clarkson S. Fisher, Jr., P.J.A.D., in the Superior Court of New Jersey, Appellate Division.

[Michael Zogby](#)



Partner-in-Charge, New Jersey Office

Barnes & Thornburg LLP

Michael C. Zogby is co-chair of the firm's Trial and Global Disputes practice, managing partner of the New Jersey office, member of the firm's Management Committee, and co-chair of the Middle East & North Africa services team. He is a trial lawyer who represents pharmaceutical, medical device, biotechnology, life sciences, manufacturing, and chemical companies in high stakes complex commercial, product liability, mass and toxic tort, IP, class actions, and environmental litigation. He also counsels clients in the U.S., Asia, and Europe on privacy, security, eDiscovery, and cross-border issues. Zogby is nationally recognized by Chambers, Legal 500, and LMG Life Sciences. Mike is vice president of the Trial Attorneys of America, treasurer/secretary of the Trial Attorneys of New Jersey,

on the board of directors of the Product Liability Advisory Council (PLAC), and a member of the Lawyers Advisory Committee for the District of New Jersey. He also serves as a faculty member for the National Trial Advocacy College at William & Mary Law School and the University of Virginia School of Law, and lectures on trial ethics and cybersecurity at Fordham University School of Law.