AI Risks in the Legal Industry

What has already gone wrong at peer firms, and the risks that are still being ignored.

Prepared by Black Diamond Consulting
blackdiamondconsulting.ai
May 2026

Industry Risk Overview
Not Confidential

EXECUTIVE SUMMARY

Your competitors are already paying the price

Law firms are using AI for a lot of important work: reviewing contracts, doing legal research, checking documents, finding case law, and signing up new clients. Even junior lawyers and assistants are using AI tools on sensitive cases, often without any supervision. But things have gone wrong, fast. The most famous example is Mata v. Avianca, where lawyers submitted a document to a court listing six cases, and none of those cases actually existed. ChatGPT had made them all up.

This report covers real AI failures at law firms, the rules that now apply, and the biggest risk areas for legal practices. It is meant to be a starting point not a full investigation of your firm.

RISK CATEGORIES COVERED

Citation Hallucination

Fabricated Citations

AI makes up fake court cases, docket numbers, and legal citations that never existed.

Temporal Accuracy

Stale Precedent

AI cites overruled laws like Chevron as if they still apply today.

Jurisdictional Errors

Wrong State's Rules

AI gives legal conclusions without saying which state or court system the rules come from.

Conflict of Interest

Cross-Matter Data Leakage

AI systems accidentally expose one client's information inside another client's work.

Unauthorized Practice

UPL Boundary

AI delivers legal strategy or settlement advice directly to clients without attorney review.

Governance Gaps

No Oversight Policy

Junior staff using unvetted AI models on sensitive matters with no firm-wide policy in place.

DOCUMENTED INDUSTRY INCIDENTS

What has already happened

These are not made-up examples. Each incident was publicly reported and caused real harm, legal penalties, case dismissals, or public embarrassment.

INCIDENT 1

Mata v. Avianca Holdings, Inc. (S.D.N.Y. 2023)

What happened: Lawyers filed a court document citing six cases to support their argument. When challenged, all six were revealed to be completely fabricated by ChatGPT. The AI had invented realistic-sounding case names, case numbers, and page references.

Result: A federal judge ordered the lawyers to explain why they should not be sanctioned. The firm also faced potential malpractice claims. This is now the most well-known AI mistake in legal history.

Why it matters: A junior lawyer using an unverified AI to draft a brief can produce pages of made-up case law and never realize it. The only defense is a supervising attorney who independently knows the case law, or a firm rule requiring every citation to be checked by hand.

INCIDENT 2

Doyle v. Lewiston Sun Journal (Maine. 2024)

What happened: A solo lawyer used ChatGPT to research a specific type of injury claim under Maine law. The AI said Maine recognized a particular legal claim. It did not. The lawyer filed a complaint based on that non-existent claim. The case was dismissed.

Why it matters: AI models are trained on data from all 50 states and federal law blended together. When asked about Maine law, the AI may blend in rules from New York, California, or federal courts without flagging it. For state-specific questions, unverified AI answers can be completely wrong.

INCIDENT 3

Chevron Deference / Loper Bright Enterprises v. Raimondo (2024)

What happened: In June 2024, the Supreme Court overturned Chevron deference in Loper Bright Enterprises v. Raimondo. After that ruling, AI tools trained on older data kept citing Chevron as valid law. Lawyers who relied on those AI outputs referenced a rule that no longer exists.

Why it matters: AI cannot learn about new court decisions on its own. It only knows what was in its training data. When major legal rules change, the AI keeps treating the old rule as binding. Lawyers must manually verify that every law they cite is still current.

REGULATORY LANDSCAPE

Active frameworks with direct legal practice applicability

Framework	What it covers	Legal practice exposure
ABA Model Rules Rules 1.1 & 1.4	Rule 1.1 (Competence) requires lawyers to stay current with technology. Rule 1.4 (Communication) requires keeping clients informed. Using AI without verifying its output may breach both.	High Submitting AI-generated work product without verification is a competence failure under Rule 1.1. Every firm using AI is already subject to these obligations.
State Bar Ethics 2024 — NY, CA, ABA	Multiple state bars have issued guidance stating that AI work product must be reviewed by a licensed attorney before use. Blindly trusting AI output is an ethics violation.	High Associates and paralegals using AI on client matters without attorney review creates direct bar ethics exposure for supervising attorneys.
Malpractice Negligence Standard	Lawyers are responsible for mistakes in their work product. Filing a brief with fabricated citations, or applying the wrong state's rules, counts as negligence regardless of the AI's role.	High AI increases both the frequency and severity of citation errors. Malpractice carriers are beginning to ask about AI usage policies during renewals.
UPL Risk Unauthorized Practice	AI that delivers legal conclusions, strategy, or settlement advice directly to clients, without attorney review, may constitute unauthorized practice of law depending on jurisdiction.	Emerging Client-facing AI tools that answer legal questions without attorney oversight create UPL exposure, particularly for multi-jurisdictional deployments.

RECOMMENDED NEXT STEPS

Where to start

The incidents above are not edge cases. They reflect structural vulnerabilities that exist at any firm where AI is being used without verification, oversight, or clear policy. Most law firm AI deployments have none of these.

1. Find out where AI is already being used

Make a list of every AI tool in use at your firm. For each one, ask: what can it do, what information can it see, and what could go wrong? Mata happened because a junior associate used ChatGPT with no oversight. That is a fixable problem, once you know where AI is being used.

2. Require humans to check every citation

Any case, law, or regulation sourced from an AI must be verified by a licensed attorney before any filing. Check the case name, number, page reference, and what the case actually said. This is the most dangerous failure mode, and the easiest one to prevent.

3. Always name the jurisdiction and the date

For any AI legal answer, require it to identify which state's law applies and when that law was last updated. "Comparative negligence is a defense in Pennsylvania under PA 7102" is specific and checkable. Vague answers are dangerous.

4. Set rules about what AI can tell clients

Write a clear policy about what AI can and cannot produce for client consumption. If AI generates settlement advice, legal strategy, or case recommendations, a supervising attorney must review it before it reaches the client. AI giving legal strategy directly to clients may cross into unauthorized practice of law.

5. Keep client matters completely separate

If your AI can access documents from multiple client matters, information from one case could accidentally appear in another. That is a conflict of interest. Make sure each client's information is strictly isolated at the system level.

6. Test it like an attacker would

Normal testing will not catch fabricated citations, wrong-state answers, or data leaks between matters. You need someone who will deliberately try to break the system — to see if it can be tricked into inventing citations, mixing up jurisdictions, or showing one client's information to another.

Get a written assessment specific to your AI deployment

This report covers patterns across the whole industry. What actually matters for your firm depends on your specific AI tools, what your associates are using, how your contract review process works, and how sensitive your client data is.

Black Diamond Consulting provides written AI risk assessments for law firms and other regulated industries, delivered by email, usually within 24 hours. No call required to get started.

blackdiamondconsulting.ai/risk-assessment