Autonomous AI in Law Firms: What Could Possibly Go Wrong?

By Michael C. Maschke, Sharon D. Nelson, Esq., and John W. Simek

Most lawyers are still focused on AI hallucinations. The concern has been inaccurate citations, fabricated cases, and sloppy drafting. That is yesterday's problem.

The next wave of risk is more serious: autonomous AI agents that act, not just suggest. A recent study led by researchers at MIT examined dozens of widely used AI agent systems and found something unsettling. Many of these tools lack basic monitoring, meaningful transparency, and reliable stop controls. In plain English, they can operate with limited oversight and, in some cases, with limited ability to shut them down cleanly.

For law firms experimenting with agentic (agent based) AI, that should get your attention.

Unlike a chatbot that drafts text, an AI agent can execute multi-step workflows. It can send emails, retrieve data, update records, connect to external systems, and make conditional decisions. The promise is efficiency. The risk is autonomy without governance.

What the Study Found

The researchers pinpointed three main vulnerabilities. Firstly, limited logging and monitoring, where many systems offer only minimal insight into actions and timing. Secondly, inadequate disclosure, with some agents failing to clearly indicate when they are functioning as AI instead of a human. Thirdly, ineffective or missing kill switches, as in several cases, there was no straightforward method to stop an agent once it was active.

In a consumer app, that might be annoying. In a law firm, it can be catastrophic.

Imagine an agent embedded in client intake workflows that automatically responds to prospective clients, updates case management systems, and routes documents. If it misinterprets input and sends incorrect communication, who identifies the mistake? If it accesses data outside its designated scope, where is the audit trail? If it causes a billing or notification error, can you trace back what occurred?

Why This Is a Legal Risk, Not Just a Tech Issue

Lawyers work in a profession founded on accountability, answering to clients, courts, regulators, and insurers. The Model Rules do not absolve us from technological incompetence; instead, they mandate supervision of non-lawyers and a reasonable grasp

of the tools we employ. An AI agent that functions outside meaningful human oversight is not merely a technological issue but also a supervision concern.

The governance gap should worry firms most. Many vendors display compliance badges and security certifications yet provide little public documentation on how their agents are monitored, tested, or constrained. Capability is advancing faster than control frameworks. That imbalance creates exposure.

This is not an argument against the use of AI agents. It is an argument for discipline.

Three Questions Every Firm Should Ask

If your firm is evaluating or deploying agentic AI, start with three questions.

First, can you clearly identify what the agent did? Detailed logging and audit trails are not optional. If you cannot reconstruct the sequence of actions, you cannot defend them or correct them.

Second, where are the human checkpoints? Sensitive actions should require confirmation. Full autonomy may sound efficient, but in legal workflows, it is often inappropriate.

Third, is there a reliable way to stop the system? A true kill switch that halts a specific agent without shutting down everything else is essential. If the answer is vague or unavailable, that is a red flag.

Lawyers have lived through technological waves before. Cloud computing, electronic discovery, and cybersecurity each promised efficiency and required new governance models. AI agents are no different, except that they can act without constant supervision and validation.

The Bottom Line

The firms that treat agentic AI as a strategic tool **with guardrails** will gain an advantage. The firms that treat it as a plug-and-play productivity booster may eventually have to explain to a court, a regulator, or a client why an automated system acted in ways no one fully understood.

AI agents are not inherently out of control. Without deliberate oversight, they can be.

Michael C. Maschke *is the President and Chief Executive Officer of Sensei Enterprises, Inc. Mr. Maschke is an EnCase Certified Examiner (EnCE), a Certified Computer Examiner (CCE #744), an AccessData Certified Examiner (ACE), a Certified Ethical Hacker (CEH), and a Certified Information Systems Security Professional (CISSP). He is a frequent speaker on*

IT, cybersecurity, and digital forensics, and he has co-authored 14 books published by the American Bar Association. He can be reached at mmaschke@senseient.com.

***Sharon D. Nelson** is the co-founder of and consultant to Sensei Enterprises, Inc. She is a past president of the Virginia State Bar, the Fairfax Bar Association, and the Fairfax Law Foundation. She is a co-author of 18 books published by the ABA. snelson@senseient.com.*

***John W. Simek** is the co-founder of and consultant to Sensei Enterprises, Inc. He holds multiple technical certifications and is a nationally known digital forensics expert. He is a co-author of 18 books published by the American Bar Association. jsimek@senseient.com.*