DIGITAL MOUNTAIN®

WINTER 2026 E-NEWSLETTER

At Digital Mountain, we assist our clients with their electronic discovery, digital forensics, cybersecurity, and data analytics needs. For this E-Newsletter, we discuss our predictions for 2026's most impactful trends on eDiscovery and cybersecurity. We address AI technology innovation, security risks and AI lawsuits coming online.

Three Critical Cybersecurity Threats for 2026: Agentic AI, AI-Powered Attacks, and Data Poisoning

As 2026 unfolds, artificial intelligence plays a dualistic role in cybersecurity, simultaneously the most powerful defensive tool and the most dangerous attack vector. While security teams deploy AI to detect threats and automate responses, adversaries weaponize the same technology to launch sophisticated attacks that evolve faster than human defenders can respond. Three critical threats are converging: autonomous AI agents operating in networks, AI-powered attacks, and data poisoning that corrupts the Large Language Models (LLMs). The rules of AI security are changing in 2026, and we have three predictions on the pivot happening.

Prediction #1: Security for Agentic AI Becomes Critical

Organizations are deploying AI agents without necessarily understanding the unprecedented autonomy with which these systems can operate. These agents operate under legitimate credentials, move laterally through networks, and make decisions that appear authorized because, technically, they are. But how do you secure something that acts without human oversight or intervention? Agentic AI breaks the assumptions of traditional security that assumes human oversight.

Insider threat detection systems fail when the "insider" is an AI agent operating within normal parameters. A compromised agent looks identical to a legitimate one, accessing the same systems with the same credentials for purposes that seem reasonable until you realize the agent has been instructed to exfiltrate data or escalate privileges. Audit trails become labyrinths obscuring which actions were explicitly authorized, which were decisions within scope, and which represent malicious deviation. When multiple agents communicate with each other, lateral movement risks multiply. Suddenly, the AI agent you deployed last month might be your biggest vulnerability today.

For 2026, we predict a rise in strict permission boundaries for AI agents, treating them as high-risk entities, recognizing that speed without security is reckless. Most critically, organizations will begin establishing "agent identity management" that is distinct from user identity management, precisely because agents operate differently than people.

Prediction #2: AI-Powered Threats: Deepfakes and Autonomous Attacks

The attacks are coming from inside the house, or appearing to, thanks to deepfakes now indistinguishable from reality. AI generates convincing phishing emails at scale, each one customized to its target based on scraped social media, previous interactions, and psychological profiling. Voice cloning enables CEO fraud where the attacker sounds exactly like your boss. Video deepfakes make video calls unreliable. Beyond social engineering, AI conducts reconnaissance, discovers vulnerabilities, and launches attacks that adapt in real-time based on defender responses. They learn while doing.

2026 will see a reshaping of cybersecurity fundamentals. Traditional authentication methods that rely on voice verification or video calls will no longer provide reliable identity confirmation. Security awareness training that focuses on old school phishing giveaways loses effectiveness when AI-generated emails are grammatically perfect, contextually appropriate, and emotionally manipulative. Last decade’s IT tools will be too slow against adaptive attacks that evolve faster than humans can synthesize. The sheer volume and sophistication of AI-powered attacks will overwhelm human defenders.

Security measures will move to methods AI can't easily replicate, such as context-based verification, out-of-band confirmation, and behavioral patterns over time rather than single-point validation. AI-powered detection systems will become the norm to fight AI attacks. Humans will still remain the weakest link in the security chain, but training will adapt to new procedures such as calling a known number before making that wire transfer.

Prediction #3: Data Poisoning and Authenticity - Corrupting LLM Models

The most insidious threat may be the quietest: attackers deliberately feeding false data into AI training sets to corrupt LLMs. Data poisoning subtly degrades model reliability over time, causing AI to make confident but incorrect decisions. Adversaries inject misinformation into datasets that organizations use to train proprietary models. Public LLMs face coordinated poisoning campaigns where bad actors contribute corrupted data. The corruption looks like legitimate data initially, making detection difficult until the damage manifests in flawed outputs.

Identifying when and how poisoning occurred becomes a forensic nightmare, especially when corruption was gradual across multiple data ingestion cycles. Retraining models after detecting corruption is expensive and time-consuming, potentially requiring organizations to rebuild from scratch. Business processes dependent on AI outputs lose credibility when stakeholders can't trust model reliability. Competitors could engage in sabotage through targeted model degradation, and you might not realize it until strategic decisions have been made based on faulty AI analysis.

Rigorous data validation and source tracking for all AI training data will become a priority in 2026. Every little thing the model does is magic—until it's been poisoned, then it's a security crisis.

How Will We Make It Through 2026?

When artificial intelligence is both a target and a weapon, it is critical that we adapt quickly. The same technology that promises to revolutionize cybersecurity also provides cybercriminals with new capabilities. Organizations that treat AI security as an afterthought will find themselves outmatched. Those building AI-specific security strategies now will have critical advantages when—not if—they face these threats.

2026 will separate prepared organizations from those caught flat-footed. Traditional cybersecurity must evolve to address AI-specific risks that didn't exist in previous threat models. The time to build AI security capabilities is today, before a breach forces reactive scrambling. We built AI on digital infrastructure, and now we must secure it against AI-powered threats.

Please direct questions and inquiries about electronic discovery, digital forensics, cybersecurity, and data analytics to info@digitalmountain.com.

UPCOMING INDUSTRY EVENTS

NETDILIGENCE CYBER RISK SUMMIT
Miami Beach, FL: February 9-11, 2026

PLANET CYBER SEC CONFERENCE
Orange County, CA: February 11, 2026

MAGNET VIRTUAL SUMMIT 2026
February 23-27, 2026

13TH ANNUAL UF LAW E-DISCOVERY CONFERENCE
Gainesville, FL: February 24-26, 2026

THE SEDONA CONFERENCE WORKING GROUP 6 ANNUAL MEETING 2026
Atlanta, GA: November 13, 2025

Click here to see more upcoming events and links.

Digital Mountain, Inc. Founder and CEO, Julie Lewis, will be presenting at various upcoming industry events. Please send requests for speaker or panel participation for her to marketing@digitalmountain.com.

DIGITAL MOUNTAIN, INC.
4633 Old Ironsides Drive, Suite 401
Santa Clara, CA 95054
866.DIG.DOCS

www.digitalmountain.com

Contact us today!

FOLLOW US AT: