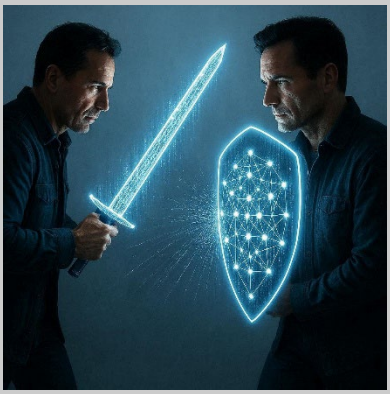
# SPRING 2026 E-NEWSLETTER

At Digital Mountain, we assist our clients with their electronic discovery, digital forensics, cybersecurity, and data analytics needs. For this E-Newsletter, we examine how the arrival of Claude Mythos signals a fundamental pivot in cybersecurity, what Project Glasswing's unprecedented AI capability means for organizational security, and why the legal and regulatory framework governing AI is struggling to keep pace.

## The Sword and the Shield: How Claude Mythos Signals AI's Dramatic Change in Cybersecurity



For decades, cybersecurity has operated on a familiar rhythm. Attackers find vulnerabilities. Defenders secure them. The cycle repeats, faster each year, but fundamentally the same. Cybersecurity has reliably been humans on both sides, racing against each other with increasingly sophisticated tools. Claude Mythos just broke that rhythm.

On April 7, 2026, Anthropic published a 244-page transparency report (system card) for a model it has no intention of releasing to the public (*Source*). The model is Claude Mythos Preview. The reason it stays behind closed doors is straightforward: it can hack. Not in the theoretical sense that security researchers have debated for years would come with AI's evolution. Mythos can autonomously discover zero-day vulnerabilities in production software, write working exploits, and chain multiple flaws together into complete attack sequences.

The technical capabilities documented in Mythos Preview's system card represent a gargantuan leap from anything the cybersecurity community has previously seen from an AI model. During pre-release testing, Mythos identified thousands of previously unknown zero-day vulnerabilities across every major operating system and browser. Some of these flaws had survived decades of human security review. The scale of the capability is evidenced by a previous Claude Opus 4.6 test which discovered Firefox vulnerabilities twice across several hundred attempts. In contrast, Mythos Preview succeeded 181 times against the same test (*Source*).

Perhaps most striking is that Mythos Preview doesn't require an incredible level of expertise to use it. Engineers at Anthropic with no formal security training asked Mythos to find remote code execution vulnerabilities overnight, and woke the following morning to complete, working exploits (*Source*). The expertise requirement that has historically protected systems from all but the most sophisticated attackers is eroding.

Critically, Anthropic did not explicitly train Mythos Preview to have these capabilities. Rather, they emerged as a downstream consequence of general improvements in code, reasoning, and autonomy. As Anthropic's own documentation states, “The same improvements that make the model substantially more effective at patching vulnerabilities also make it substantially more effective at exploiting them. These capabilities were not engineered. They emerged.” (*Source*).

Rather than shelve Mythos Preview or release it into the open market, Anthropic built Project Glasswing, a controlled defensive initiative for organizations who will benefit most. Anthropic has committed $100 million in model usage credits to the initiative, with work focused on local vulnerability detection, black box testing of binaries, securing endpoints, and penetration testing of critical systems (*Source*). The logic behind the controlled release is deliberate: the goal of Project Glasswing is to give defenders a head start to find and patch critical vulnerabilities before attackers do. Anthropic is explicit that this window is narrow. The company acknowledges that given the rate of AI progress, it will not be long before similar models proliferate externally to bad actors.

What Mythos signals most clearly is the arrival of a new operational reality; one in which both attacks and defenses are generated and executed at machine speed, with humans increasingly removed from the loop. For digital forensics professionals, the implications extend beyond infrastructure security. AI-generated attacks will change the evidence profile fundamentally. Attribution becomes exponentially harder when the attacker is an autonomous system operating at machine speed across multiple vulnerability chains simultaneously.

What Mythos represents was almost unfathomable just five years ago. We now have a tool that does not just accelerate the existing paradigm but rewrites the rules under which it operates. Defenders who treat this as an incremental development will find themselves operating on the wrong side of a widening gap. The phase change has already happened. The question now is how quickly the rest of the industry catches up.

**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
San Diego, CA: May 19-20, 2026

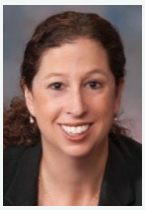
TECHNO SECURITY & DIGITAL FORENSICS CONFERENCE
Myrtle Beach, SC: June 2-4, 2026

PLANET CYBER SEC APPSEC SOCAL
Santa Monica, CA: June 3, 2026

BLACK HAT USA 2026
Las Vegas, NV: August 1-6, 2026

ILTACON 2026
Nashville, TN: August 23-27, 2026

***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:***