AI, Mythos, and the cybersecurity inflection point

July 2026

Tapestry Networks recently convened a group of board directors and executives from companies headquartered around the world for a series of discussions focused on the introduction of Anthropic's Mythos and Fable models, and the broad impact that AI is having on cybersecurity. The conversations highlighted how the release of Mythos marked what many in the security community now regard as an inflection point, with AI emerging as both a new source of cyber risk exposure and an indispensable defensive capability. The discussions revealed that the right response will not be technical alone, but organizational—rooted in governance, accountability, and operating discipline. Boards and executives are now grappling with how advanced AI models are fundamentally reshaping cyber risk, and what they must do in the narrow window before adversaries possess the same capabilities.

Mythos is both a little hype and a very real and fundamental signal

The first task for any senior business leader is to separate signal from noise. While participants broadly agreed that Mythos represents a turning point, some emphasized that much of the public narrative around Mythos confuses what the model actually does with what people fear it might do. The technology is genuinely transformational, but it is not a magic weapon that compromises any target on command. Participants offered several perspectives on the prevailing threat narrative:

- **Mythos matters, but mainly as a signal of broader change.** One participant described Mythos as partially *"the best marketing of the year,"* partially a real step change in capability, and absolutely a useful wake-up call for boards and executive leadership teams. Another participant called it *"both hype and truth,"* suggesting that the restricted release likely reflected compute limits as much as caution, while acknowledging that the underlying capability gain was real. The model appeared to be roughly 20% better than its predecessor at identifying and attacking vulnerabilities. However, that participant was more skeptical of exaggerated public narratives that imply the model alone can autonomously *"hack a company,"* stating, "*A friend told me that his organization pointed Mythos at their website,*

and Mythos was able to compromise it in 10 minutes. Mythos does not do that. That's not how Mythos works." Mythos requires additional software and infrastructure to function, and, even in its preview form, retained significant guardrails and safety rules to restrict malicious use. The consensus was to focus less on Mythos as a specific tool and more on the broader trajectory it represents. *"The trend line is clear,"* an executive said, *"This was a step change that has caused us to react in a way that we probably should have been reacting six months ago. And if it weren't Mythos, something else would make us react this way in another six months."*

- **Mythos is best understood as a defensive advantage, not merely an offensive weapon.** The dominant media framing casts Mythos as a destructive force, but the practitioners closest to it described it as a tool that could enable organizations to harden systems faster than attackers can exploit them. A participant explained, "*Mythos is so advanced that it's finding the needles in the haystack, which is really good from a defense posture. You can actually harden your system significantly more than the attackers can gain."*
- **Other models are available and may be more dangerous than Mythos.** Participants noted that models built explicitly for cyber-attacks were already available in early summer and may be more effective in enabling bad actors than Mythos. Restricting one model does little if equivalent capability exists elsewhere. *"There are other models that are far better at offensive security, things like the OpenAI 5.5 cyber model that does not have guardrails on it,"* a participant warned. In addition to OpenAI models, the best open-source models are expected to reach comparable cybersecurity capability by the end of 2026, according to a participant: *"I think the right assumption is by the end of 2026, this will be a capability that is broadly available. And, frankly, I think it will actually be two or three months earlier than that, although it's a reasonable assumption to just plan for the end of the year."*

Guardrails and the politics of access are a new challenge

The release and almost immediate shutdown of Fable exposed how little consensus exists about who controls access to these models and on what basis. Participants surfaced several unresolved questions about guardrails and model governance:

- **Model-level guardrails are real but probabilistic, and organizations must build their own.** Because large language models are non-deterministic, no guardrail can be fully vetted. A participant candidly shared, *"You can never really fully test an LLM to understand what the guardrails actually are."* One leader noted that their firm's internal monitoring alerted the organization to an employee's potential misuse of AI tools, illustrating that companies can and must establish safeguards and visibility independent of the guardrails built into the underlying models.
- **AI is becoming increasingly politicized, resulting in unpredictable access.** Asked why

the Trump administration first intervened only with the release of Fable, a participant suggested the answer may lie partly in the relative power of the models and partly in commercial and political relationships beyond public view. But this does not necessarily discredit the principle of restriction: *"With these powerful models, we should know who's running them. They shouldn't just be wild and open to everybody. There should be some governance around it. And then there's some accountability across the board."* Subsequent restrictions placed on powerful OpenAI models suggest that the debate around restrictions will not end any time soon, and that access granted today may not be available tomorrow.

- **Restrictions heighten existing concerns around sovereign technology.** The Mythos saga has exacerbated concerns many senior leaders outside of the US share about their organizations' over-reliance on US tech providers. A participant noted that European organizations appear even more keen to build sovereign models that are independent of US tech and the unpredictability of US policy. *"Europe is trying to build their own independent models. They're worried about the administration and whether they'll be able to have access to the models in Europe, but I do think that the race for AI is super powerful and whoever has the best AI will probably still be the dominant market,"* said a participant.

The current cybersecurity environment exposes new risks

Even if Mythos is partly AI hype, the risks the model exposes are nonetheless significant. Two related exposures emerged in the discussions: risks in the software supply chain and the collapse of the traditional model for managing vulnerabilities. A third exposure, the new risks created by adopting AI itself, runs beneath both.

Open-source and software supply chain risks are a central vulnerability

Open-source dependencies and compromised binaries—pre-built or third-party software components that may be embedded into a company's tech stack—are one of the most pressing risks, and one that AI tools now make dramatically easier to exploit. Virtually all modern software rests on a foundation of open-source libraries, and the practices around acquiring and updating those libraries have not sufficiently evolved. Participants identified several dimensions of software supply chain risk:

- **Open source is ubiquitous.** *"Your organizations are almost certainly built on a foundation of open source that Mythos will be run against to find vulnerabilities,"* said a participant. One executive affirmed this statement saying, *"We completed a scan of all of our internet-facing applications, and 100% of them had some level of open source in them, whether it was code that used open-source or third-party software that we include in the code we build that uses open source."*
- **Compromised binaries injected into open source are the most dangerous emerging**

attack. One participant described how attackers are already using AI to insert malicious functionality into open-source libraries. *"It's a lot easier to do this today than it was six months ago,"* the participant said, calling it "*the single most vital and important problem."* When a company deploys the compromised library, the attacker gains remote access that the organization cannot see. The scale of the problem could be immense, considering that large companies often depend on 25,000 to 50,000 open-source libraries, many of which are redundant. A participant argued that disciplined organizations could reduce that number to around 5,000, allowing developers to focus and maintain a more finite set.

- **Vibe coding and automatic updates amplify the exposure.** The risk is compounded by development practices that automatically pull the latest versions of libraries without sufficient scrutiny. A participant warned that *"where vibe coding is going today is it's just downloading random things. And those random things that are part of the build of vibe coding, that many people do not know about, are having these remote-control capabilities built into them that the attackers can access. It's happening across the internet."*

Vulnerability management is no longer keeping pace

The window between vulnerability disclosure and exploitability is collapsing, undermining legacy response cycles and rendering them no longer fit for purpose. Existing processes—a vulnerability is reported, it's named, it receives a severity ranking, companies update their scanners, then they patch—depended on having days or months before an exploit appeared. *"All of those steps are going to be under enormous stress and possibly collapse. That supply chain is almost certainly going to break and never be reconstituted,"* said an executive. Participants described how traditional approaches are changing:

- **The time from disclosure to exploit has shrunk to hours.** What once took years, then months, has now compressed dramatically. *"We're seeing a massive shortening of the time it takes to write code to exploit a vulnerability. And that's an area that Mythos and other models excel at. It used to be years and then months, but earlier in 2026 it went down to about a day. It's probably around about an hour now,"* shared a participant.
- **New capabilities mean constant updating, not periodic scanning and patching.** Rather than maintaining lists of known vulnerabilities and checking exposure, organizations must move toward continuously refreshing software whether or not a specific vulnerability is known. *"That is extremely difficult to do in a way that doesn't create operational problems,"* shared a participant. As one participant noted, the best technology companies are already releasing new features to customers within hours of writing them.
- **Memory-safe programming languages and foundational controls remain vital.** Amid the hype and hysteria surrounding advanced models, participants stressed that the basics are even more relevant. An executive suggested that companies using memory-safe languages, which are designed to eliminate entire categories of common coding errors,

were significantly more secure than software built using older programming techniques. Another participant was blunt about the consequences of neglecting other well-established fundamentals: *"If you're not using phishing-resistant multifactor authentication, you're not washing your binaries, you're not patching your systems, you don't stand a chance today."*

Adopting AI creates its own new risks

While adopting advanced AI may be the key for organizations to defend themselves against increasingly sophisticated bad actors, broader deployment of the technology also introduces its own set of risks. Participants highlighted a couple requiring senior attention:

- **AI-assisted workflows create new attack opportunities.** Employees across industries are utilizing AI in an ever-increasing number of ways for their day-to-day work, creating new opportunities for bad actors to infiltrate systems. For example, as one participant noted, employees are delegating email triage to AI, exposing them to attackers who embed instructions invisible to a human reader directly in messages. While models may have guardrails against following malicious instructions, researchers have found that rephrasing such instructions as poetry could fool nearly every model. The participant warned, *"That's the kind of thing that is so far outside of the normal security thinking paradigm that we're going to be learning about and addressing these attacks for years."*
- **AI elevates the insider threat from clumsy to sophisticated.** A participant warned that more sophisticated AI tools turn *"your average hacker into a ninja hacker. The problem is they also do that for your insider threat."* Many current insider-threat controls are predicated on the idea of making things more difficult for a malicious insider and ensuring they must make multiple attempts before succeeding. Coding tools collapse that gap. A participant warned that these tools let an insider *"skip the attempts and go straight to the sophisticated evasion of your control, and we may not see it. You have to be using these AI tools, but the risk of an insider threat, or a more sophisticated insider threat, is way higher as a result."*

Mythos requires a fundamental response from organizations and their boards

Participants emphasized that the evolving threat environment requires a reorientation of the entire technology organization, deliberate trade-offs, and board-level attention to culture and accountability.

Embrace the paradox: use AI to defend against AI

There is no path forward that avoids using AI to combat AI. The tools that create new risks are also the only means of keeping pace with adversaries who will soon wield equivalent capability. As one leader framed it, *"You're being forced into a corner where you have to use AI to address the risks AI has created, even though the AI itself will create other risks,"* because *"the risk of*

not using AI is going to be higher." Participants discussed several approaches organizations can take to utilize AI to strengthen cyber defenses:

- **Build a robust cyber testing environment now, with or without Mythos.** Waiting for access to currently restricted models is the wrong strategy. A participant urged, *"All organizations should be building a testing harness, whether you have Mythos or not. Mythos is great for very complex reasons, but the cost of it is not sustainable. Even with access, you can't afford to run Mythos full time. So, using different models in conjunction with Mythos is really the plan forward. Even if Mythos is better, it doesn't mean the other models aren't good."*
- **Run multiple models against one another.** Because any given model can make mistakes, participants recommended having independent models check the work of others. One participant said they ask any given model to challenge their own findings and run seven additional models alongside it, treating cross-validation as a basic discipline of secure AI deployment.
- **Prioritize the most exposed and highest value assets first.** *"You're eventually going to have to apply it to everything,"* observed a participant, *"but you have to start with the most important things."* Strong authentication should be at the top of the list followed by more secure use of AI software development tools, better segmentation or barriers that prevent attackers from moving freely across systems if they gain a foothold, and finally more resilient backups. Ultimately, according to a participant, *"there's no real revolution in thinking about what you do outside of vulnerability management to reduce your risk. It's the same stuff we've been talking about for years, but the urgency is just dramatically increased."*

Curate third parties toward a smaller, trusted core

Third-party risk emerged as both a top concern and perhaps the hardest to manage. Participants advocated reducing the number of vendors to a smaller set of well-resourced, trusted providers and noted that much vendor redundancy is accidental rather than strategic, with different business lines buying largely the same products from different vendors. The goal is to find the balance between using the smallest possible number of suppliers and maintaining an appropriate level of resilience and business support. Participants shared emerging principles for managing vendor risk:

- **Utilize AI to reduce and curate the vendor base.** One leader proposed using AI to recreate capabilities that were previously bought from third parties, eliminating the risk of onboarding a less-secure provider: *"I think trust becomes more relevant in this world. And so, if you're a vendor and you solve a problem, but I don't have a ton of trust in you, and I could write my own solution, and I have trust in my internal processes, maybe I do write my own solution."* Another participant agreed with this approach, saying, *"Anytime you onboard a third party, you're taking on risk. So, why would we allow a third-party chatbot to connect*

to our environment when it's something we can probably build in two or three weeks?"

- **Treat smaller, fast-growing vendors with particular caution.** Small and midsized companies, even if generating significant revenue, often lack mature controls. *"They have a million different existential risks that they're worried about, not just cybersecurity. So, cybersecurity is very rarely the top existential risk for them, and they invest accordingly,"* observed a participant. Another shared similar concerns: *"How many security people can you put in a $100 million company? And is that number going to be representative of your organization and the level of sophistication that you're doing? My favorite question to ask third parties is, how many security people are in your organization? If you come back with three, I'm not going to do business with you."* On the other hand, an executive noted, *"I don't think there's any rule that says, go with the old established vendors. Because plenty of them have been coasting on archaic software delivery processes for a long time and are going to be in trouble too."* Regardless of size, cybersecurity should weigh more heavily in vendor selection than perhaps it has in the past.

Treat culture and accountability as the real problem

Security failures are often cultural as much as technical, and boards too often focus on technical details when the underlying issue is actually accountability. Participants explored a board-level approach grounded in metrics and ownership to tackle this challenge:

- **Use a small set of metrics to surface accountability gaps.** A participant shared that ideal reporting to the board should include three simple measures: 95% of known vulnerabilities should be remediated within the service-level agreement, 95% of defined controls should be operating as intended, and all cyber incidents should be detected and contained within 24 hours. They noted that when these metrics slip, the cause is rarely technical. *"If you're not hitting those, it's probably cultural. In my experience, it wasn't the technical work that was the issue, it was always cultural."*
- **Ask who is accountable, not what the vulnerability is**. Rather than asking which systems carry which flaws, boards should ask who owns each control and why standards are not being met. One executive recalled telling a board that was fixated on operating-system versions, *"You're asking the wrong question. Who is accountable for delivering these vulnerabilities? It doesn't matter if they're Windows 95 or Windows 2000, who's responsible for that component? I think I would go there and look at accountability. So, you don't put people on the spot. And that will probably also tell you if there's good alignment within management."*
- **Demand evidence of action, not intention, on AI.** Boards should expect to see concrete movement from their management teams, from testing harnesses in development, multiple models in use, and higher-quality findings at greater volume. Mere plans are insufficient given the timeline.

Questions boards should be asking about AI and cybersecurity

- ? How are we controlling the open-source libraries and binaries we depend on, and how do we know they are secure, particularly where automatic updates and AI-assisted coding are in use?
- ? Given that the window from vulnerability disclosure to exploit has collapsed to hours, how is our technology organization responding?
- ? What share of our vulnerabilities are remediated within our service-level agreements, what share of our defined controls are operating as intended, and how quickly do we contain incidents—and where these slip, who is accountable?
- ? How are we curating our third-party vendors in this new environment, and does cybersecurity carry appropriate weight in how we select and concentrate our critical suppliers?
- ? As we deploy AI agents and coding tools, what guardrails are in place to mitigate the new risks AI creates?