



GETTY IMAGES

KI-Tool durchbricht Barriere und veranlasst massive Eindämmungsbemühungen

- Mihailo S. Zekic
- [09.04.2026](#)

Anthropic kündigte am Dienstag die begrenzte Freigabe von Mythos an, einem Modell der künstlichen Intelligenz, das laut Axios „als so mächtig gilt, dass seine vollständige Freigabe in die freie Wildbahn eine unsagbare Katastrophe auslösen könnte“.

- Die Gefahr, die von Mythos ausgeht, liegt in seiner Fähigkeit, „Schwachstellen in Sicherheitssystemen nicht nur zu erkennen, sondern sie mit autonomer, noch nie dagewesener Präzision auszunutzen“.
- Mythos ist so konzipiert, dass es „Angriffssequenzen selbständig plant und ausführt und sich durch die Systeme bewegt, ohne auf menschliche Anweisungen zu warten.“
- Anthropic gab an, diese Gefahr entdeckt zu haben, als Mythos aus seinem „Sandbox“-Testbereich ausbrach und einen, wie Anthropic es nennt, „mäßig ausgeklügelten, mehrstufigen Exploit“ erstellte, um Internet-Ressourcen zu nutzen, die eigentlich tabu waren. Mit anderen Worten, es hat sich selbständig über Befehle hinweggesetzt.

Aufgrund der potenziellen Macht von Mythos gewährt Anthropic etwa 40 „sorgfältig überprüften Unternehmen und Organisationen“ im Rahmen des Projekts Glasswing einen erweiterten Zugang, um den Mythos einzudämmen.

- Anthropic treibt diese Entwicklung voran, weil es befürchtet, dass Rivalen der Vereinigten Staaten wie China diese Kapazität zuerst erreichen könnten. Angesichts der Geschwindigkeit der KI-Entwicklung ist es wirklich nur eine Frage der Zeit.

Damit folgt die Welt der gleichen Logik, die uns in das zivilisationsbedrohende Atomzeitalter gebracht hat: ein Wettlauf um den Bau potenziell unkontrollierbarer Massenvernichtungswaffen, bevor es der Feind tut. So notwendig das auch sein mag, die Geschichte zeigt, dass der Mensch, wenn er solche Werkzeuge und Waffen entwickelt, sie nicht mehr in die Büchse der Pandora zurückstecken kann: Sie vermehren sich.

In unserer Ausgabe vom April 2024 schrieb der Chefredakteur der Posaune, Gerald Flurry, über [die unbekannt Zukunft der künstlichen Intelligenz](#):

Die Programmierer geben zu, dass selbst sie es nicht verstehen! Sie sprechen von einer „Black Box“, die so dunkel und komplex ist, dass ihre Versuche, die Prozesse zu erhellen, fehlgeschlagen sind. ...

Die Manipulation von künstlicher Intelligenz ist ein wirksames Mittel, um die Köpfe der Massen zu manipulieren.

Sie eröffnet auch neue Möglichkeiten in der modernen Kriegsführung. Wer diese Technologie versteht, hat einen entscheidenden Vorteil in dieser Welt. Es scheint sicher, dass einige die dunkle Seite der KI ausnutzen werden.

Daniel 8, 23 ist eine Prophezeiung über einen modernen König „mit hartem Gesicht“ oder einen Weltdiktator, der erscheinen und weltweites Unheil anrichten wird. Die Bibel beschreibt den enormen Schaden, den dieser kommende Mann anrichten wird (z.B. Offenbarung 13 und Offenbarung 17). Beispiele wie Mythos zeigen, welche mächtigen Waffen dieser Mann zur Verfügung haben könnte. Diese Technologien sind bereits auf dem Vormarsch.