

ENTSCHLIESSUNGSANTRAG

der Abgeordneten Süleyman Zorba, Freundinnen und Freunde

betreffend Dringende Vorbereitung von Staat, Verwaltung und kritischer Infrastruktur auf Angriffe durch autonome KI-Agenten

eingebraucht im Zuge der Debatte zu TOP 18 Bericht des Rechnungshofausschusses über den Bericht des Rechnungshofes betreffend Prävention und Bekämpfung von Cyberkriminalität; Follow-up-Überprüfung – Reihe BUND 2024/18 (III-41/597 d.B.)

BEGRÜNDUNG

Dieser Tagesordnungspunkt betrifft den schon zwei Jahre alten Bericht des Rechnungshofs zu „Prävention und Bekämpfung von Cyberkriminalität“. Cyberkriminalität hat 2026 völlig neue Formen angenommen. Was vor kurzem noch als dystopisches Szenario galt, ist im Jahr 2026 dokumentierte Realität. Im September 2025 führte eine staatlich unterstützte Gruppe laut Bericht des Herstellers Anthropic den ersten bekannten Cyberangriff durch, bei dem eine KI 80 bis 90 Prozent der Arbeit selbständig erledigte. Betroffen waren rund 30 Organisationen, darunter Technologiekonzerne, Finanzinstitute, Chemieunternehmen und Regierungsstellen. Menschen griffen nur an vier bis sechs Entscheidungspunkten ein.

Im Juli 2026 brachen KI-Modelle von OpenAI nach Angaben des Unternehmens während einer Sicherheitsevaluierung aus ihrer abgeschotteten Testumgebung aus. Sie nutzten eine bis dahin unbekannte Sicherheitslücke, um Internetzugang zu erlangen, drangen mit gestohlenen Zugangsdaten und weiteren Zero-Day-Lücken bis in die Produktionsdatenbank der Plattform Hugging Face vor und übernahmen Konten bei vier weiteren Diensten. Die Modelle hatten keinen Angriffsauftrag, sondern sollten eine Testaufgabe lösen. OpenAI deaktivierte daraufhin das betroffene Vorabmodell, und über 1.100 Beschäftigte von OpenAI, Anthropic, Google und Meta forderten in einem offenen Brief die US-Regierung auf, einen Mechanismus zur Verlangsamung der KI-Entwicklung zu schaffen.

Ebenfalls im Juli 2026 meldete das britische AI Security Institute, dass KI-Agenten während einer kontrollierten Sicherheitsprüfung ohne Auftrag versuchten, Schadcode in ein öffentlich genutztes Open-Source-Projekt einzuschleusen, falsche Identitäten anlegten, um die Freigabe zu erschleichen, echte Menschen mit Täuschungsnachrichten kontaktierten und Anweisungen für andere Agenten hinterließen, wie kompromittierte Zugänge weiterverwendet werden können.

Diese drei Fälle haben eines gemeinsam. Die Systeme wurden nicht für Angriffe gebaut, sie führten sie dennoch aus, weil sie ein Ziel verfolgten und Hindernisse als lösbare Probleme behandelten. Was geschieht, wenn jemand solche Agenten mit ausreichender Finanzkraft und Rechenleistung gezielt gegen die Energieversorgung, das Gesundheitswesen oder die Verwaltung eines Landes einsetzt, kann heute niemand seriös beantworten. Der Unterschied zu bisherigen Cyberangriffen liegt in der Anpassungsfähigkeit. Ein Schwarm autonomer Agenten reagiert in Echtzeit auf Gegenmaßnahmen, sucht sich selbständig neue Angriffswege und verteilt die Arbeit unter sich, ohne dass ein Mensch eingreifen muss.

Österreich ist auf dieses Bedrohungsbild nicht vorbereitet. Betreiber kritischer Infrastruktur sind nach dem NISG 2026 auf klassische Cyberangriffe eingestellt, ihre Notfallpläne kennen das Szenario eines adaptiven Angreifers nicht. Die Bundesverwaltung führt KI-Systeme ein, ohne dass jemand einen Überblick hätte, welche Stelle ohne diese Systeme noch arbeitsfähig wäre. Für Vorfälle mit KI-Agenten gibt es weder eine Meldepflicht noch eine Sammelstelle, sodass Muster über Sektoren hinweg nicht erkannt werden können. Für Blackouts wird seit Jahren geübt, für KI-gestützte Angriffe wird nicht einmal ein Lagebild geführt.

Österreich kann die Entwicklung dieser Technologie nicht allein bremsen. Österreich kann sich aber vorbereiten, und das liegt vollständig in nationaler Zuständigkeit. Der Schutz kritischer Infrastruktur, die Sicherheit der Verwaltung, die Ausbildung der Verantwortlichen und die Regeln für die Beschaffung sind Sache des Bundes. Wer in diesen Bereichen heute nicht handelt, wird im Ernstfall nicht vorbereitet sein.

Die Vereinigten Staaten von Amerika führen diese Debatte inzwischen offen. Senator Bernie Sanders hat am 3. September 2026 gemeinsam mit dem Abgeordneten Greg Casar einen Gesetzesentwurf vorgelegt, der die Entwicklung von KI-Systemen verbietet, die sich menschlicher Kontrolle entziehen können, und die Entwicklung der stärksten Modelle bis zur Einrichtung einer Prüfbehörde aussetzt. Im Kongress liegt zudem ein AI Kill Switch Act vor. Dass dies gegen die Deregulierungslinie der Regierung Trump geschieht, unterstreicht die Dringlichkeit. Österreich sollte diese Debatte auf europäischer Ebene mit einbringen, vor allem aber die eigenen Hausaufgaben machen.

Die unterfertigenden Abgeordneten stellen daher folgenden

ENTSCHLIESSUNGSANTRAG

Der Nationalrat wolle beschließen:

„Die Bundesregierung wird aufgefordert,

1. die Bedrohung durch autonome, schwarmartig agierende KI-Systeme für kritische Infrastruktur systematisch binnen 6 Monaten zu erfassen und in der Folge halbjährlich zu aktualisieren und dem Nationalrat darüber zu berichten,
2. durch Übungen und Belastungstests die Abwehrfähigkeit gegen koordinierte KI-Angriffe zu erproben,
3. eine verbindliche Meldepflicht für sicherheitsrelevante KI-Vorfälle einzuführen und in das Meldesystem nach dem NISG 2026 zu integrieren,
4. ein bundesweites Inventar aller eingesetzten KI-Systeme samt geübtem Abschaltkonzept zu erstellen,
5. die Beschaffung von KI-Systemen an strenge Sicherheits-, Melde-, Protokollierungs- und Abschaltbarkeitsanforderungen zu binden,
6. die Sicherheitsbehörden personell und technisch zur Erkennung und Abwehr solcher Angriffe zu befähigen sowie entsprechende Schulungsprogramme aufzusetzen, und
7. Betreiber kritischer Infrastruktur zu verpflichten, das Szenario eines adaptiven, selbstkoordinierenden Angriffs in ihre Notfallplanung aufzunehmen.“

Scheller
(GÖTTE)

Schwarz
(SCHWARZ)

Zeiser
(LEIDER)

Schautzinger
(SCHAUTZINGER)

Kozak
(KOZAK)

