KI in der Cybersicherheit: Gekommen, um zu bleiben

Künstliche Intelligenz ist ein Dauerbrenner und verändert die Cybersecurity-Landschaft rasant. Die Experten der Sophos X-Ops skizzieren die wichtigsten Trends und Sicherheitsherausforderungen, mit denen Unternehmen im Jahr 2026 in Sachen KI rechnen müssen und zeigen auf, welche neuen Risiken wie auch Chancen sich daraus für Unternehmen ergeben.

1. KI-gestütztes Programmieren: Der neue blinde Fleck der Sicherheit
KI-Programmierplattformen wie Replit, Lovable, GitHub Copilot und Cursor senken die Einstiegshürden für die Softwareentwicklung drastisch und befeuern einen Boom neuer webbasierter Start-ups. Diese Demokratisierung der Entwicklung wird schnelle Innovation ermöglichen, setzt Anwender jedoch zugleich erheblichen Sicherheitsrisiken aus. Vielen Start-ups fehlt die Erfahrung im Aufbau sicherer Architekturen, wodurch Anwendungen durch schwache Authentifizierung, falsch konfigurierte APIs und mangelhafte Datenverarbeitung verwundbar werden. Software wird also einfacher denn je zu entwickeln sein – ist aber nicht zwangsläufig sicherer.
Allerdings gibt es auch einen Lichtblick. KI-gestütztes Programmieren wird ebenfalls Innovationen bei der Entdeckung und Behebung von Schwachstellen vorantreiben, insbesondere im Open-Source-Ökosystem. Ein bemerkenswertes Beispiel ist CodeMender vom DeepMind-Team von Google: Während Programmierassistenten eine deutlich schnellere, aber auch fehleranfälligere Softwareentwicklung ermöglichen, erlauben LLMs die proaktive Identifikation von Sicherheitslücken in großem Maßstab, was dazu beitragen wird, dieses Risiko auszugleichen.

2. Die eigentliche Angriffsfläche: Die KI-Anwendung
Wahrscheinlich werden wir innerhalb des nächsten Jahres größere Sicherheitsvorfälle durch Prompt-Injection-Angriffe erleben. Über Jahre hinweg haben Sicherheitsteams daran gearbeitet, ihre Internet-Angriffsfläche zu verkleinern, im Bewusstsein, dass jede exponierte Komponente das Risiko erhöht. Firewalls, VPNs und ZTNA zielten allesamt darauf ab, diese Angriffsfläche zu reduzieren. Nun haben wir fast über Nacht eine neue geschaffen: schnell bereitgestellte KI-Anwendungen. Viele davon sind aus dem Internet erreichbar, häufig ohne Authentifizierung, und mit Daten verbunden, die viele Unternehmen als sensibel oder vertraulich einstufen würden. Noch besorgniserregender ist, dass diesen Anwendungen oftmals die Befugnis eingeräumt wird, im Namen der Organisation zu handeln. Das hohe Tempo der KI-Einführung bringt enorme Effizienzgewinne – doch wenn Unternehmen nicht innehalten und diese Risiken bewerten, öffnen sie Schwachstellen erneut, die sie über Jahrzehnte hinweg mühsam geschlossen haben.

3. Die Risikophase von Multi-Agenten-Systemen
Multi-Agenten-Systeme bewegen sich rasant von der Forschung in die praktische Anwendung und bringen dabei neue Sicherheitsherausforderungen mit sich. Wir werden 2026 vermutlich die erste Welle von „Living-off-the-Land“-Angriffen sehen, die auf Agenten abzielen, denen ungesicherte Zugriffe auf interne Systeme gewährt wurden. Als Reaktion darauf werden Organisationen strengere Berechtigungssysteme einführen, die klar festlegen, auf welche Ressourcen jeder Agent zugreifen darf und wie Daten zwischen ihnen fließen dürfen.

4. Blue Teams liegen im KI-Wettrüsten vorn
KI bietet aktuell Verteidigern einen Vorteil, der sich auch in den nächsten Monaten positiv auswirken wird. Blue Teams können frei auf hochmoderne LLMs zurückgreifen, während Angreifer zunehmend eingeschränkt werden. Große Anbieter wie OpenAI und Anthropic

entfernen aktiv Bedrohungsakteure aus ihren Ökosystemen und begrenzen damit deren Fähigkeit, Schadcode oder Exploits zu erzeugen. Gleichzeitig nutzen Verteidiger dieselben Tools, um stärkere Automatisierungs-, Erkennungs- und Reaktionsfähigkeiten aufzubauen. Zum ersten Mal seit Jahren könnten die Guten tatsächlich leicht im Vorteil sein. Allerdings wird sich diese Lücke schnell schließen, sobald leistungsfähigere Open-Source-Modelle günstiger und einfacher bereitzustellen sind. Blue Teams können sich über ihren Vorsprung freuen, doch es gibt keinen Raum für Nachlässigkeit. Jetzt ist der Moment, den Vorteil auszuspielen, aggressiv zu bleiben und die Oberhand zu behalten.

5. Der nächste Insider ist die KI
Organisationen wetteifern darum, LLMs und Agenten zur Effizienzsteigerung einzuführen. Indem sie diese Tools mit riesigen Mengen an Unternehmensdaten füttern, schaffen sie eine neue Klasse von Insider-Bedrohungen. Wenn diese Daten abfließen, wer trägt die Verantwortung? Ist die KI ein „Mitarbeiter“ – und wer haftet, wenn sie außer Kontrolle gerät, kompromittiert oder falsch konfiguriert wird? Diesen Fragen müssen sich die Verantwortlichen im Unternehmen im Vorfeld stellen und eine dedizierte Sicherheitsstrategie zur Nutzung solcher Tools aufsetzen.

Social Media von Sophos für die Presse
Wir haben speziell für Sie als Journalist*in unsere Social-Media-Kanäle angepasst und aufgebaut. Hier tauschen wir uns gerne mit Ihnen aus. Wir bieten Ihnen Statements, Beiträge und Meinungen zu aktuellen Themen und natürlich den direkten Kontakt zu den Sophos Security-Spezialisten.

Folgen Sie uns auf und

LinkedIn: https://www.linkedin.com/groups/9054356/
X/Twitter: @sophos_info

Pressekontakt:

Sophos
Jörg Schindler, Senior PR-Manager EMEA Central
joerg.schindler@sophos.com, +49-721-25516-263

TC Communications
Arno Lücht, +49-8081-954619
Thilo Christ, +49-8081-954617
Ulrike Masztalerz, +49-30-55248198
Ariane Wendt +49-172-4536839
sophos@tc-communications.de