# Hackerintelligenz und KI-Kompetenz

## Die Zukunft von Penetrationstests


Severin Quell, Matthias Szymansky

**Das Wettrüsten zwischen Angreifern und Verteidigern im Netz geht in eine neue Runde. Hacker setzen immer häufiger auf den Einsatz künstlicher Intelligenz (KI). Security-Analysten schlagen zunehmend mit vergleichbaren Waffen zurück. Severin Quell, Director IT Security Consulting und Matthias Szymansky, Director IT Consulting, beide bei der Infodas GmbH, stellen KI-gestützte Angriffsmethoden vor und denken über die Zukunft von Penetrationstests nach.**

Severin Quell ist Director IT Security Consulting und Matthias Szymansky Director IT Consulting bei der Infodas GmbH in Berlin

Die Basis jeder Softwareanwendung und jedes Endgerätes ist von Menschen geschaffener Softwarecode, der potenziell Sicherheitslücken beinhaltet und damit Einfallstore für Cyberattacken bietet. Weil das Ausnutzen von Schwachstellen arbeitsintensiv sein kann, machen sich Hacker immer häufiger schlagkräftigere Methoden und Tools zunutze, um IT-Schutzmaßnahmen effizient und intelligent zu unterwandern, zum Beispiel mithilfe künstlicher Intelligenz. KI beflügelt die Dynamik von Bedrohungen und Gefährdungen und kann mögliche schädliche Auswirkungen durch automatisierte oder autonome Prozesse weiter potenzieren.

Captchas etwa sind ein gutes Beispiel dafür, wie intelligent KI inzwischen sein kann. Sinn der Captchas war es bisher, die Interaktion mit echten Menschen von der mit Bots zu unterscheiden. Genau das wird immer schwieriger. Mittlerweile existieren KI-Anwendungen, die fähig sind, Captchas wie ein Mensch zu lösen.

### Gefahr durch KI-Angriffe

Mit Unterstützung von künstlicher Intelligenz lassen sich inzwischen auch Intrusion-Detection-Systeme (IDS), Data Leak Prevention (DLP) und Cloud Access Security Broker (CASB) unterwandern, herkömmlicher Netzverkehr oder das Verhalten eines authentifizierten Nutzers oder seines Endgerätes simulieren – und nicht zuletzt eigene Aktivitäten verschleiern.

*Mit Unterstützung von künstlicher Intelligenz lassen sich mittlerweile Intrusion-Detection-Systeme (IDS), Data Leak Prevention (DLP) und Cloud Access Security Broker (CASB) unterwandern (Foto: Darwin Lager, Pixabay)*

Auch die Angriffsintensität und Erfolgsquote von Spear-Phishing und Advanced Persistent Threats (APT) lassen sich steigern durch KI, der Aufwand für den Angreifer verringert sich. KI kann für Hacker Aufgaben übernehmen, für die bislang noch menschliche Intelligenz notwendig war. Während APTs oft groß angelegte Angriffe sind, die beispielsweise kritische Infrastrukturen und Behörden auf dem Radar haben und dazu viele Einzeltechniken verwenden, ist das Spear-Phishing eine konkrete Angriffstechnik, die darauf abzielt, Personen zu manipulieren und die Schwachstelle Mensch auszunutzen.

KI ist in diesem Zusammenhang sehr hilfreich, um soziale Netze, Business-Plattformen, Onlineshops und Foren zu durchforsten und künstliche Beziehungs-

netze aufzubauen. Durch KI-gesteuerte Chat-Bots sendet der Hacker Kontaktanfragen an potenzielle Opfer. Nach der Annahme der Kontaktanfrage sammelt er weitere Informationen und wertet sie aus. Persönliche Anschreiben mit fingierten Absenderadressen und einem auf Basis von KI trainierten Schreibstil können so täuschend echt wirken, dass die Menschen immer wieder auf Anhänge mit Schadsoftware oder Links zu gefälschten Webseiten hereinfallen.

## Auffinden von Schwachstellen

Zur Champions League zählt die Aushebelung biometrischer Authentifizierungsverfahren, die künftig durch Fälschung oder Manipulation angreifbar werden könnten. Bei sogenannten Deepfakes, einer Wortkombination von Deep Learning und Fake, lassen sich Audio-, Video- oder Bilddateien durch KI mit neuen Inhalten besetzen und täuschend echt manipulieren.

„AI Augmented Systems“ sind Systeme, die durch KI (engl. AI – Artificial Intelligence) verstärkt und in ihrer Wirkung noch unterstützt werden. Eine Technik, die sich gut mit KI anreichern lässt, ist das Fuzzing. Für Hacker ist Fuzzing eine bekannte Methode, um Schwachstellen in Systemen zu finden. Dieser Prozess, der bislang manuell durchgeführt werden musste, ist wegen der großen Menge an möglichen Eingaben sehr aufwendig. Mit KI lässt sich die Arbeit automatisieren, potenzielle Schwachstellen werden ressourcenschonender und rascher gefunden und lassen sich damit auch schneller ausnutzen. Daher nennt man diese Methode auch AI Fuzzing.

## Angriff auf die KI

Daneben gibt es Methoden, die direkt KI-Systeme angreifen. Eine verdeckte Manipulation von Trainingsdaten durch Hinzufügen oder Entfernen von Daten wird als Machine Learning Poisoning bezeichnet. Der einfachste Typ ist die sogenannte Availability Attack. Hierbei werden so viele „schlechte” Daten hinzugefügt, dass das Modell nutzlos wird.

Gefährlicher sind Integrity Attacks, das heißt, Manipulationen des Bias gegenüber bestimmten Daten, während das Modell hinsichtlich anderer Daten nicht verändert wird – eine Art Backdoor, die es erlaubt, schädliche Datentypen als harmlos durchgehen zu lassen.

Werden solche KIs in Sicherheitssystemen eingesetzt, haben sie genau diese eingebauten Backdoors, die Angreifer später unbemerkt ausnutzen können. Aus diesem Grund gewinnt das Supply-Chain-Management gerade bei der KI-Entwicklung

*Security-Analysten von morgen müssen über Hackerintelligenz und KI-Kompetenz verfügen*

immer mehr an Bedeutung. Ob es sich um Sicherheitssysteme wie IDS handelt oder um Systeme, die hochsensible Daten verarbeiten, es muss gewährleistet werden, dass der Entwicklungsprozess nachvollziehbar bleibt und Manipulationen möglichst ausgeschlossen sind.

## Wirksame IT-Sicherheitsstrategien

Klar ist: Die hier aufgeführten Beispiele sind nur die Spitze des Eisbergs. Umso wichtiger ist es, Cyberkriminelle mit ihren eigenen Waffen zu schlagen. KI kann Cybersecurity-Experten von Routineprozessen entlasten, damit sie sich besser auf ihre Kernaufgaben konzentrieren können. Sie ist eine Unterstützung mit Blick auf Echtzeitanalysen und schnelle kontextbasierte Informationsauswertung. Mittlerweile sind die ersten Security-Information- and Event-Management-Systeme (SIEM) mit KI auf dem Markt. Auch Data-Loss-Prevention-Lösungen (DLP) und Netzüberwachung mit KI-Funktionen in Echtzeit erhöhen die Sicherheit des Informationsverbundes. Allerdings werden sie ihre Schutzwirkung nur so lange behalten, wie die Systeme schneller lernen als die der Angreifer.

Darüber hinaus ist es wahrscheinlich, dass wir bald in der Lage sein werden, mit neuen KI-Werkzeugen zur Erkennung und Abwehr von Cyberangriffen effektiv zu arbeiten. Ein Konzept für ein solches Werkzeug beschreibt ein Frühwarnsystem für APT-Angriffe. APT-Angriffe funktionieren üblicherweise in aufeinander aufbauenden kleinen und sehr schwer zu entdeckenden Schritten. Gelingt es also, einen dieser Aktionen zu vereiteln, bricht der gesamte Angriff in sich zusammen. Hierzu muss eine Entdeckung von APT in Echtzeit durch gleichzeitige Überwachung aller Zugänge zu Unknown Domains, maliziöse DNS (Domain Name System) und URL sowie Malware realisiert werden. Machine Learning wird eingesetzt, um harmlose sowie schadhafte Domains zu klassifizieren. Dabei gilt auch hier: Je größer die Menge an qualitativ hochwertigen Trainingsdaten, desto präziser ist die Vorhersage.

*www.infodas.de*

### Menschliche Intelligenz versus KI

KI ist längst eine wichtige Größe, sowohl auf Seiten der Angreifer als auch bei der Abwehr von Cyberattacken. Ohne menschliche Intelligenz – etwa in Form von Penetrationstests – wird es auf lange Sicht deshalb keinen Schutz vor Cyberkriminellen geben. Penetrationstests werden im Rahmen wirksamer IT-Sicherheitskonzepte deshalb eine noch wichtigere Rolle spielen als bisher.

Die Verringerung von False Positives, also das Erkennen von falsch-positiven Warnungen sowie die Absicherung und Überwachung der KI durch menschliche Intelligenz wird künftig immer wichtiger.