VON **BARBARA STEININGER**

Schweizermesser für Cyberbetrug

KÜNSTLICHE INTELLIGENZ beflügelt nicht nur Kreative, sondern auch Kriminelle im Cyberspace. Sie gehen mit ungekannter technischer Raffinesse gegen Unternehmen und Privatpersonen vor.

Ein Freitagnachmittag im April. Das Smartphone eines bekannten Wiener Unternehmers läutet, am Apparat die weinerliche Stimme der 16-jährigen Tochter, die in der Schweiz zur Schule geht: Sie habe einen großen Blödsinn gemacht, der zwei Menschen das Leben gekostet hätte. Danach übernimmt ein Polizist das Gespräch und will den geschockten Vater anleiten, auf welche Bank er die 100.000 Euro Kaution überweisen könne, um die Tochter vor der Untersuchungshaft zu bewahren. Zwei Details ließen die Eltern stutzig werden. Die zufällig anwesende Mutter, eine Juristin, wusste, welche Abläufe bei solchen Delikten tatsächlich in Gang kommen – und der aus Genf anrufende Polizist sprach Englisch.

Die Eltern erstatteten Anzeige bei der Polizei und wiesen ihre Tochter, die sie zum Zeitpunkt des Vorfalls nicht erreichen konnten, weil diese einen Test schrieb, zu erhöhter Vorsicht an. Was ihnen besonders in die Glieder gefahren war: „Es ist beunruhigend, wie täuschend echt die Stimme unserer Tochter war."

Das Paar wäre fast Opfer eines „Enkeltricks" geworden, wie Ermittler die Betrugsmasche nennen, bei denen Opfern suggeriert wird, ein naher Verwandter befände sich in ernster Gefahr: Sicherheitsexperten sprechen hier von Social Engineering, wo menschliche Gefühle wie Angst ausgenutzt werden. Die Stimme der Tochter war wohl, so vermutet die Polizei, von ihrem TikTok-Account abgenommen und präpariert worden, Kontaktdaten und mutmaßliches Vermögen des Vaters ließen sich im Netz relativ leicht recherchieren.

ChatGPT und ähnliche KI-Programme zeigen einmal mehr das Dual-Use-Dilemma bahnbrechender Technologiesprünge. Denn mit den smarten Text- und Bildassistenten munitionieren sich nicht nur Wissensarbeiter auf, sondern auch Kriminelle. Und das lässt Cybercrime in einem bislang ungekannten Ausmaß eskalieren.

„Künstliche Intelligenz ändert nichts an den Grundmustern, aber erweitert die Möglichkeiten, das Potenzial im Bereich von Cybercrime auszuschöpfen und es auf ein neues Level zu heben", sagt Robert Herscovici, Geschäftsführer der Sicherheitsfirma TCSS. „Phishingmails kommen mit exzellenter Textierung daher, Profile in Netzwerken lassen sich dank perfekter Bild- und Textprogramme in beliebiger Zahl herstellen und ausspielen", sagt der Experte (siehe auch Kästen).

Diese KI-Werkzeuge verändern die Spielregeln nachhaltig, Cyberangriffe lassen sich dadurch noch besser skalieren, so Herscovici: „Es wird viel mehr und viel bessere Angriffe geben, weil die Angreifer mit weniger Personaleinsatz die Qualität massiv steigern werden können."

Mittels ChatGPT lassen sich sogar Sicherheitslücken im Quellcode von Software recherchieren und mit entsprechenden Programmierkenntnissen natürlich auch ausnutzen. ChatGPT entlastet also nicht nur die ehrlichen Programmierer, ChatGPT entlastet auch die Bösen.

FRÜHE WARNUNGEN. Die jüngsten rasanten Entwicklungen im Machine Learning lassen in der Sicherheitsindustrie die Alarmglocken schrillen. Europol, Trend Micro und UNICRI (das Forschungsinstitut für Kriminalität der UNO) warnten bereits Ende 2020 in einem großen Report vor den Gefahren des kriminellen Missbrauchs der neuen ►

„Es wird viel mehr und viel bessere Angriffe geben, weil die Angreifer mit weniger Personaleinsatz die Qualität massiv steigern können."
ROBERT HERSCOVICI, TRUSTED CYBER SECURITY

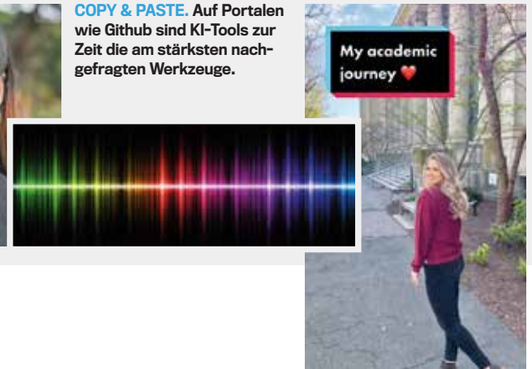
FOTOS: ISTOCKPHOTO (2), LUKAS ILGNER, SHUTTERSTOCK, BEIGESTELLT

Zwischen künstlerischer Freiheit und krimineller Energie

Einzigartige biometrische Merkmale wie Gesicht und Stimme lassen sich mit simplen Programmen so einfach und günstig wie noch nie manipulieren, vervielfältigen und für klandestine Zwecke missbrauchen.

MANIPULATION. Zwölf Minuten reichen, um eine KI-Stimme zu trainieren – mit solchen Versprechen lockt die Industrie. Sie verspricht nicht zu viel: Was früher nur ein Tontechniker machen konnte, geht heute mit einer simplen App, die Laien bedienen können. Ausgelotet werden die Grenzen nicht nur von Kreativen, sondern auch von Cyberkriminellen. Eine Vielzahl günstiger oder kostenloser KI-Programme erlaubt das Manipulieren audiovisueller Inhalte in verblüffender Qualität. Gesichter werden für Laien nicht erkennbar ausgeschnitten und mit anderen überblendet. Tonspuren von TikTok, YouTube und Co. werden geklont und mit neuen Inhalten aufbereitet.

COPY & PASTE. Auf Portalen wie Github sind KI-Tools zur Zeit die am stärksten nachgefragten Werkzeuge.

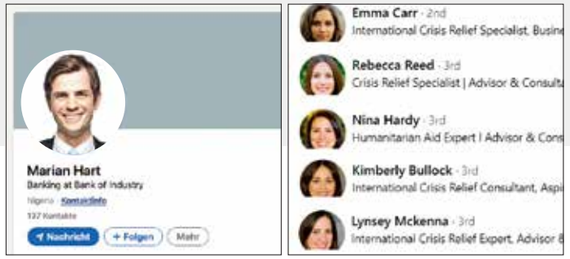


2018 fanden US-Forscher heraus, dass künstliche generierte Gesichter in Videos nicht normal blinzeln, weil die Algorithmen meist mit statischen Gesichtern trainiert werden. Ein Manko, das rasch ausgebessert wurde; Lichteinfall und Mimik werden heute perfekt animiert. Die großen Netzwerkbetreiber wie Facebook oder Google versuchen seit Jahren, mithilfe von KI Deepfake-Inhalte zu identifizieren und einzudämmen. Auch der politische Druck auf die Netzwerke wird nun erhöht: China hat unlängst ein Deepfake-Gesetz erlassen, und die EU versucht, das heiße Eisen mit dem jüngsten KI-Act zu regeln (siehe Seite 33).

LinkedIn: zwischen Fake-Profilen und gefährlicher Mitteilsamkeit

Fake-Accounts sind ein Problem auf allen großen Plattformen: Beim populären Karrierenetzwerk LinkedIn wird es für Unternehmen besonders gefährlich. Hinter professionellen Fassaden tarnen sich Spione und Betrüger.

FAKE FRIENDS. Das populäre Karrierenetzwerk ist für Kriminelle eine einträgliche Informationsquelle. Massenhaft künstlich generierte Profile suchen Kontakt (Bild l.). Gern genutzt werden prestigeträchtige Arbeitgebermarken wie Apple. Allein 2022 warf Microsoft 250.000 falsche Apple-Profile aus dem Netzwerk (Bild rechts). Genutzt werden solche Profile, um Opfer bzw. Unternehmen auszuspionieren und Erpressungen vorzubereiten. Vorsicht, wenn Unbekannte zu beruflichen Leistungen gratulieren. Kontakte überprüfen (etwa Google-Bildersuche), bevor sie ins eigene Netzwerk gelassen werden, Sicherheitseinstellungen nutzen und klare Richtlinien für die Firmenkommunikation ausgeben, die von Führung, Marketing/PR und Security formuliert werden. Viele Nutzer mit zu vielen Details ihrer beruflichen und privaten Vita sind unnötig angreifbar. Öffentliche Daten werden vom Netzwerk „gescraped" und in Hackerforen im großen Stil zum Verkauf angeboten.



► Technologie: „Künstliche Intelligenz als Dienstleistung legt die Einstiegsbarriere noch tiefer und erlaubt technischen Laien, elaborierte Attacken zu fahren und Sicherheitsbarrieren anzugreifen." Die Experten beschrieben schon damals das immense Gefahrenpotenzial von mittels KI gefälschten, täuschend echt wirkenden Fotos sowie Audio- und Videoaufnahmen, genannt Deepfakes.

Kaspersky hat dieser Tage eine Untersuchung veröffentlicht, die nahelegt, dass die Nachfrage nach Deepfake-Videos das Angebot im Darknet übersteigt. Ab 300 Dollar kann dort bei „Profis" die Erstellung beauftragt werden. Das Angebot reicht vom privaten Racheporno bis zur Rufschädigung öffentlicher Personen. „Die hohe Nachfrage deutet darauf hin, dass Einzelpersonen und Gruppen mit krimineller Energie bereit sind, erhebliche Summen für den Erwerb solcher Videos zu zahlen", warnt Kaspersky-Analyst Vladislav Tushkanov. Beispielen wie einer betrunkenen Nancy Pelosi oder einem ausrastenden Robert Habeck werden viele folgen.

Security-Experte Herscovici bereitet das schon lange Sorgen: „Mit Fake News wird professionell Stimmung gemacht und die öffentliche Meinung beeinflusst. Unternehmen bieten ganze Botfarmen als Dienstleistung an, um Wahlen oder Aktienkurse zu beeinflussen." Ein demokratiepolitisches Problem, das in seinen Dimensionen völlig unterschätzt wird. „Im deutschsprachigen Raum ist das Problembewusstsein dafür noch kaum vorhanden, in Österreich sehe ich es überhaupt nicht."

Hergestellt und trainiert werden die Falschinformationen mit öffentlich verfügbaren Sprechproben der Opfer, die Auftritten oder Social-Media-Posts entstammen und neu abgemischt werden. So werden etwa Sprachdateien von CEOs manipuliert, um Überweisungen von Töchterfirmen zu veranlassen. Besonders perfid: Selbst das Interview eines früheren Europol-Chefs wurde bereits für Erpressungen missbraucht.

BUCHHALTUNG IM VISIER. Den Enkeltrick gibt es übrigens auch in der Firmenvariante, er tritt auch in Österreich immer häufiger auf. Marc Nimmerrichter von Certitude Consulting: „Angreifer geben sich bei den Verrechnungsabteilungen als Lieferanten aus, erwirken eine Änderung der Kontoverbindungen und veranlassen Zahlungen." Die Angreifer wirken glaubhaft, beziehen sich auf aktuelle Projekte und kennen die verantwortlichen Personen. Geschäftskorrespondenz nachzuahmen, fällt ihnen leicht. „Oft reicht ihnen eine einzige Mail an den Vertrieb mit einer allgemeinen Anfrage, auf die meist höflich geantwortet wird, und schon haben sie die passenden Signaturen", so Nimmerrichter. Allein Certitude hatte in nur einem Monat vier Vorfälle zu betreuen. Die Fake-Rechnungsbeträge lagen bei über 100.000 Euro. „Bei zwei der vier Kunden waren die Betrüger erfolgreich, und das führt nun zu Rechtsstreitigkeiten zwischen Opfer und den echten Lieferanten", bedauert der Experte. Die Mitteilsamkeit vieler Unternehmen spielt den Kriminellen in die Hände. Nimmerrichter: „Wir stellen bei den Analysen fest, dass die Betrüger gezielt Informationen zu den Unternehmen zusammentragen und genau wissen, wann sie sich gezielt in die Kommunikation einschalten können."

Informationsquellen sind Pressemeldungen, Blogposts, LinkedIn-Einträge und sogar öffentliche Ausschreibungen, die automatisiert bearbeitet und editiert werden. ChatGPT ist dabei für Kriminelle ein digitales Schweizermesser, dem permanent neue Tricks beigebracht werden.

Wer dem Programm die richtigen Fragen stellt, stellt fest, dass kein moralischer Kompass einprogrammiert ist. Wie sollten sich Unternehmen auf die neuen Möglichkeiten der KI einstellen? Sicherheitsexperten haben vor allem eine Empfehlung, und die ist so alt und bewährt wie die Betrugsmaschen. „Laufende Schulungen sind die beste Prävention", so Herscovici, der empfiehlt, die eigenen Teams regelmäßig über neuesten Tricks aufzuklären und zu trainieren, damit sie in ihrer digitalen Kommunikation „immer eine gesunde Portion Misstrauen im Hinterkopf haben". T

„Wir stellen bei den Analysen fest, dass die Betrüger gezielt Informationen zu den Unternehmen zusammengetragen haben."
MARC NIMMERRICHTER
CERTITUDE CONSULTING

Der lange Marsch zur Regulierung

Seit 2021 arbeitet die EU an der Regulierung von künstlicher Intelligenz. Regelmäßig wird sie von den Entwicklungen in diesem Feld überrollt. Nun soll der Entwurf des EU-Parlaments endlich fertig werden.

GEFAHRENZONE. Mit der Datenschutzverordnung hat die Europäische Union einen internationalen Standard gesetzt. Gleiches soll nun auch mit dem AI Act gelingen, mit dem die EU neue Entwicklungen der künstlichen Intelligenz regulieren möchte. Am Donnerstag (nach Redaktionsschluss) soll die jüngste Fassung der Richtlinie noch einmal in den zuständigen Ausschüssen des Europaparlaments behandelt werden, bevor sie in den „Trilog", also in die Verhandlungsrunde mit der EU-Kommission und dem EU-Rat geht. Bis 2024 soll der AI Act stehen und 2025 in Kraft treten.

Die rasanten Fortschritte, die sich bei neuen Anwendungen wie dem Chatbot ChatGPT oder der Bilderstellungs-KI Midjourney beobachten lassen, zeigen die Dringlichkeit klarer Regeln auf. Diese Form der KI war in dem ursprünglichen AI Act noch gar nicht berücksichtigt, das hat sich nun geändert. In den USA lud US-Präsident Joe Biden die Größen der KI-Industrie zuletzt ins Weiße Haus. China, wo nach den USA die meisten KI-Innovationen entstehen, hat erste Regeln aufgesetzt. Sie fallen streng aus und nehmen sowohl die Nutzer als auch die Hersteller rechtlich in die Verantwortung.

Den bisher umfangreichsten Zugang soll jedoch die EU vorlegen. Was aber sieht der AI Act bisher vor?

MARGRETHE VESTAGER ist als EU-Wettbewerbskomissarin zuständig für die Regulierung von Big Tech. Auch für den AI Act ist sie verantwortlich: „Es gibt keine Zeit zu verlieren."

FOTOS: BEIGESTELLT (2), LUKAS ILGNER, EU/AURORE MARTIGNONI

RISIKOANALYSE. Über die Datenschutzverordnung und anderes bestehendes Recht wie das Urheberrecht bewegt sich KI auch jetzt schon nicht im rechtsfreien Raum. Der AI Act geht jedoch weiter: Abhängig davon, für wie gefährlich ein Programm eingestuft wird, sollen strengere Regeln gelten. Stellt ein System ein „unannehmbares Risiko" dar, kann es die EU verbieten. Das träfe zum Beispiel für eine KI zu, die ein Social-Credit-System aufbauen könnte, wie es China verwendet. Nach den jüngsten Änderungen sollen auch Systeme, die biometrische Daten wie Gesichter erkennen, gänzlich verboten sein – außer es gibt schwerwiegende Sicherheitsgründe.

Wird ein System hingegen als hochriskant eingestuft, weil es „ein hohes Risiko für die Gesundheit und Sicherheit oder für die Grundrechte natürlicher Personen darstellt", greifen strenge Qualitäts- und Transparenzregeln. Das betrifft Systeme, die in der kritischen Infrastruktur wie der Energieversorgung vorkommen, aber auch jene in der Strafverfolgung. Systeme, die über Schulnoten, Bewerbungen oder Versicherungen für Einzelne entscheiden können, gelten ebenfalls als hochriskant. Menschen, die etwa in einem Bewerbungsprozess auf KI treffen, sollen darüber informiert werden müssen.

Durch die jüngsten Ergänzungen wäre auch „Generative AI" wie ChatGPT oder Midjourney reguliert, allerdings nicht automatisch, sondern nur, wenn sie ein „signifikantes Risiko" darstellt. Dann sollen die Anbieter eine Zusammenfassung der Trainingsdaten vorlegen und offenlegen müssen, dass ihre Inhalte durch KI generiert wurden. Werden diese Vorgaben nicht eingehalten, sollen Strafen von bis zu zwei Prozent des Jahresumsatz fällig werden. Ein „AI Office" soll als neue Institution zudem für EU-weite Harmonisierung sorgen.

Kritiker bemängeln den Vorschlag auf mehreren Ebenen: Forscher:innen befürchten, unklare Vorgaben könnten Open-Source-KI und somit Forschung unmöglich machen. Manche Branchen warnen vor Überregulierung, andere vor unklaren Zuständigkeiten. MBA