Künstliche Intelligenz

Doppelagentin KI

Immer öfter setzen Cyberkriminelle Künstliche Intelligenz für ihre Angriffe ein. Aber auch auf der Seite der «Guten» wird KI immer wichtiger.

KI senkt die Einstiegshürden für bösartige Aktivitäten.

Noch bestimmt der Mensch, für welche Zwecke Künstliche Intelligenz eingesetzt wird.

Kameraüberwachung per KI? Problematisch!

Text **Erik Brühlmann**
Bilder **Unsplash / Aidin Geranrekab, Solen Feyissa, Markus Spiske, Claudio Schwarz, zVg**

kompakt

KI

Künstliche Intelligenz (KI) ist zurzeit in aller Munde und längst schon in vielen Bereichen des alltäglichen Lebens präsent.

Künstliche Intelligenz (KI) ist zurzeit in aller Munde und längst schon in vielen Bereichen des alltäglichen Lebens präsent. Studierenden und manchen Medien greift sie beim Schreiben von Texten unter die Arme; in Autos bietet sie Fahrerinnen und Fahrern Unterstützung auf den Strassen; Sprachassistenten und Chatbots üben sich in Kommunikationsfragen; im Hintergrund agiert sie auch bei E-Mail-Programmen, Streaming-Diensten, in Sozialen Medien – sowie bei Überwachungs- und Sicherheitsdiensten.

Altes Konzept, rasante Entwicklung

Vereinfacht gesagt, soll KI menschliche Denk- und Verhaltensweisen erkennen und imitieren und möglichst präzis auf menschliche Bedürfnisse eingehen können. Und dies alles so selbstständig wie möglich. Das Konzept ist dabei wahrlich nicht neu: Künstliche Intelligenz hat ihre Wurzeln in den 1950er-Jahren, als der britische Mathematiker und Informatiker Alan Mathison Turing in einem Artikel die Frage diskutierte, ob Maschinen denken können. Auch Machine Learning datiert in jene Zeit zurück, genauer gesagt auf den Perceptron, einen Algorithmus für maschinelles Lernen, den der US-amerikanische Psychologe und Informatiker Frank Rosenblatt 1957 entworfen hatte. «Seit etwa einem Jahrzehnt erfährt jetzt das Deep Learning einen riesigen Schub», sagt Reto Zbinden. Er ist Gründer, Inhaber und CEO von Swiss Infosec, dem führenden Schweizer Beratungs- und Ausbildungsunternehmen in den Bereichen Informationssicherheit, Datenschutz und IT-Sicherheit. Deep Learning erlaubt die Analyse und Verarbeitung komplexer Datenmuster. «Und mit den Milliarden Bildern aus dem Internet liessen sich plötzlich Systeme gezielt darauf trainieren, Dinge zu erkennen.» Trotzdem, so der Experte, kann man heute immer noch nur von schwacher KI sprechen: Sie kann Daten analysieren, Bilder erkennen, Texte übersetzen oder erstellen – aber nur in einem begrenzten Rahmen. Sie ist nicht in der Lage, menschliche Intelligenz zu imitieren oder gar zu übertreffen.

Training ist alles

KI kommt bereits in vielen Sicherheitsbereichen zum Einsatz, vom Identitäts- über den Datenschutz bis Erkennung von und Reaktion auf Vorfälle. KI ist allerdings, wie Reto Zbinden es nennt, eine Funktionalität, die auf eine Wissensbasis zurückgreift. Und die Qualität dieser Wissensbasis ist entscheidend für die Qualität der Performance des jeweiligen KI-Systems. «Hier geht es nicht nur um die Frage des Datenschutzes, die im Zusammenhang mit KI immer wieder aufkommt», sagt Reto Zbinden, «sondern auch darum zu wissen, dass nicht mit Daten trainiert wird, die falsche Ergebnisse liefern.» Wird die KI zum Beispiel mit rassistisch gefärbtem Grundlagenmaterial trainiert, kann sie beim Einsatz zu ganz anderen Resultaten gelangen als eine KI, die mit neutralem oder ausgewogenem Datenmaterial trainiert wurde. Reto Zbinden bemängelt denn auch, dass heute oft nicht ganz transparent ist, auf welcher Datenbasis KI-Systeme trainiert wurden. «Es gibt deshalb Forderungen, dass Trainingsdaten nachvollziehbar sein müssen, auch um Missbräuche zu vermeiden», sagt er.

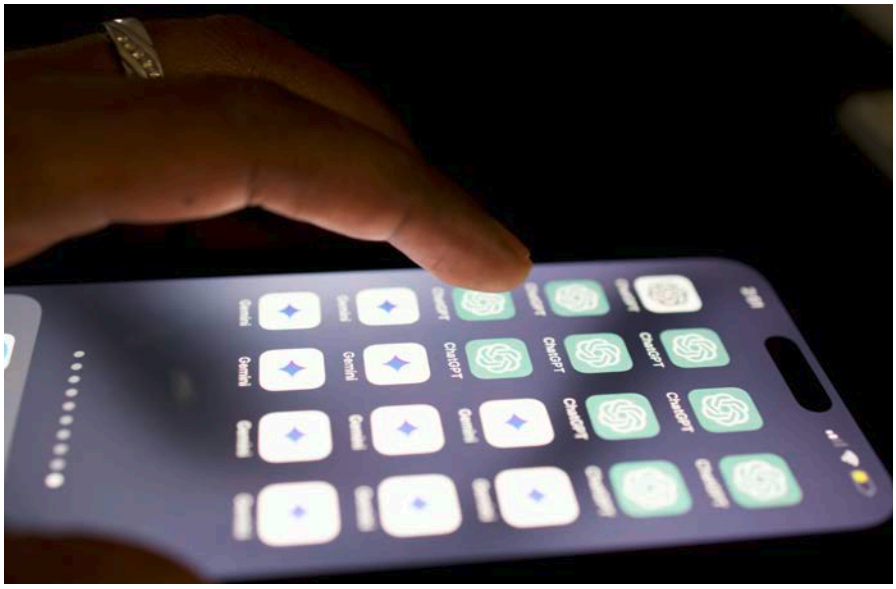
Der Mensch wird nicht abgelöst

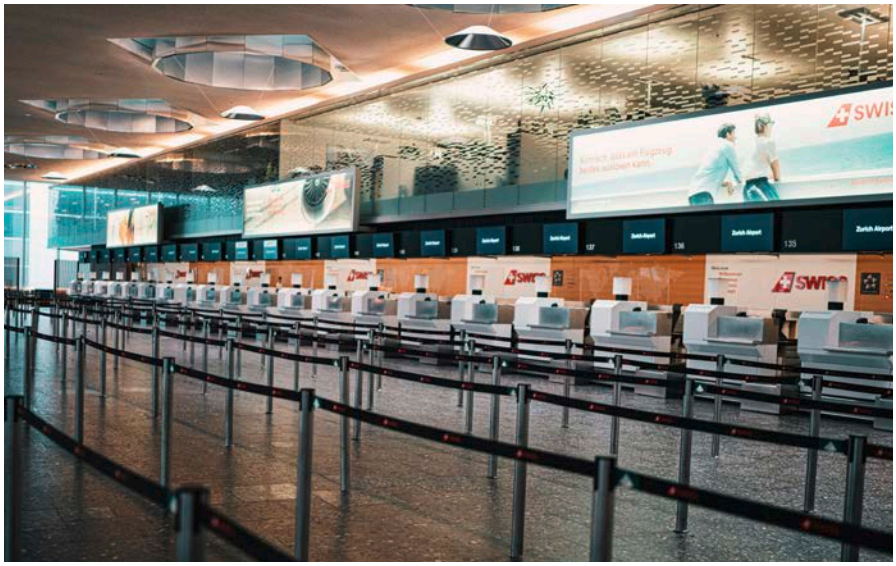
Gerade im Sicherheitsbereich ist dies eine Forderung, die auf offene Ohren stossen sollte. Ebenso wichtig ist für Reto Zbinden jedoch, KI nicht zum Selbstläufer werden zu lassen. «Ich plädiere dafür, dass der Mensch die KI-Entwicklung begleitet und kontrolliert», sagt er. «Zumindest, wenn es nicht einfach darum geht, ein Licht ein- und auszuschalten, sondern wenn Resultate aus komplexen Aufgaben besser werden sollen.» Es brauche Spezialisten, die mit KI arbeiten, um sie zu verbessern. So dient KI als Instrument zum Optimieren von Tätigkeiten. Allerdings scheint wie so oft im technischen Bereich die Entwicklung schneller voranzuschreiten, als der Mensch damit umzugehen lernt. «Noch sind viele Fragen offen, wie man KI sinnvoll begleitet und den Menschen vor KI schützt», so der Experte. Für das deutsche Fraunhofer-Institut für Kognitive Systeme IKS scheint zudem eine Software-Begleitung sinnvoll: «Da der Entscheidungsweg der KI undurchsichtig ist, kann die Sicherheit und Zuverlässigkeit der KI bisher nicht ohne Weiteres bewertet werden», heisst es auf der Website. «Diese Nachvollziehbarkeit ist aber notwendig, um Unsicherheiten der KI messbar zu machen und daraufhin dynamische Sicherheitsmechanismen zu entwerfen … Über eine adaptive, erweiterte Softwarearchitektur werden Fehler der KI abgefangen, damit die KI Menschen nicht gefährden kann.»

Schnell reagieren

Die Vorteile einer KI-Unterstützung bei Sicherheitssystemen liegen trotz ungeklärter Fragen jedoch auf der Hand. So kann KI zum Beispiel das Verhalten von Nutzern und Systemen innerhalb eines Unternehmens lernen. Abweichungen der erlernten Norm oder ungewöhnliche Aktivitäten können so schnell erkannt und angezeigt werden. Hinzu kommt, dass die Algorithmen von KI-Tools grosse Datenmengen in kurzer Zeit analysieren können. Die Reaktionszeit solcher Systeme ist deshalb kurz, wodurch Angriffe vielleicht abgewehrt werden können, bevor sie viel Schaden anrichten. Zudem wird auf diese Weise das IT-Team entlastet, da weniger manuelle Aktionen nötig werden. Wertvolle Arbeit könnten KI-gestützte Systeme auch bei der Bilderkennung auf der Suche nach bestimmten Personen liefern, zum Beispiel an Flughäfen. «Hier ist der Datenschutz jedoch ein riesiges Thema», sagt Reto Zbinden. Trotzdem können KI-Bilderkennungen die Sicherheit auf datenschützerisch unbedenklichem Weg erhöhen. Zbinden: «Sofern ein System einen vom Benutzer markierten potentiellen Laden-

Reto Zbinden, Gründer, Inhaber und CEO von Swiss Infosec.





Alle Welt beschäftigt sich mit KI – einige zum Zeitvertreib, andere aus beruflichen, wieder andere aus kriminellen Gründen.

KI wird an Flughäfen in Zukunft wohl bei der Suche nach verbotenen Gegenständen helfen.

dieb automatisch auf seinem Weg durch ein Kaufhaus verfolgt, ohne ihn dabei zu identifizieren, ist dies viel weniger bedenklich.» Und für Sicherheitsbeauftragte von unschätzbarem Wert.

Erhöhte Flughafensicherheit

Dennoch haben KI-Systeme an Flughäfen ihren Nutzen. So forscht zum Beispiel ein Team der Hochschule für Angewandte Psychologie der Fachhochschule Nordwestschweiz (FHNW) in Zusammenarbeit mit Flughäfen, Sicherheitsdienstleistern und Geräteherstellern an APIDS: Automated Prohibited Items Detection Systems. Diese sollen beim Röntgen zuverlässig verbotene Gegenstände in Passagierstücken erkennen. Vor allem Waffen und Sprengstoffe sollen dem Sicherheitspersonal so nicht mehr durch die Lappen gehen. Derzeit untersucht das Forschungsteam, wie gut APIDS schon funktionieren, wie sich die Zusammenarbeit von Menschen und Technik gestaltet und wie Sicherheitspersonal für den APIDS-Einsatz geschult werden müsste. In Zusammenarbeit mit der Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) entwickelte das FHNW-Team eigene APIDS-Algorithmen. Trainiert wurde das System mit Bildern verbotener Gegenstände aus unterschiedlichen Perspektiven. Es ist nicht unwahrscheinlich, dass APIDS in absehbarer Zeit die Sicherheit in Flughäfen und auf Flügen erheblich verbessern können und werden.

KI als Tathelfer

Doch KI kann auch zum Sicherheitsrisiko werden. Reto Zbinden berichtet von englischen Studenten, welche die von Meta und Ray-Ban neu entwickelte smarte Brille zum Erkennen und Identifizieren von Passantinnen und Passanten nutzten. Sie liessen das Videobild der Brille über einen Instagram-Livestream laufen, der wiederum von KI beobachtet wurde. Diese machte Schnappschüsse von Gesichtern und leitete diese an einen Onlinedienst, der sie mit Bildern auf Webseiten abglich – und bei Treffern Namen und meist auch persönliche Informationen lieferte. Studenten können das? Tatsächlich stellte das Deutsche Bundesamt für Sicherheit in der Informationstechnik BSI in einer Untersuchung fest, dass KI und vor allem grosse Sprachmodelle (LLMs) «die Einstiegshürden für bösartige Aktivitäten senken». KI habe bezüglich Cyber-Sicherheit einen «dualen Effekt», so das BSI: Sie ermöglicht zwar effizientere und Möglichkeiten zur Abwehr und Aufklärung von Angriffen. Gleichzeitig erhöht sie aber auch die Qualität und Reichweite von Cyberangriffen. «Es ist durchaus denkbar, dass ein KI-Modell bereits beim Entwickler manipuliert wird», sagt Reto Zbinden, «etwa dahingehend, dass gewisse Personengruppen aufgrund irgendwelcher Merkmale eben nicht erkannt werden sollen». Ausserdem kann man ein KI-System sogenannt vergiften: «Wenn das Trainingsmaterial manipuliert oder mit Falschinformationen versetzt wird, kann man die Effizienz des gesamten Systems verschlechtern», erklärt der Experte.

KI-Wettrüsten?

Das Wettrüsten im Bereich digitaler Sicherheit, das einst mit Viren und Anti-Viren-Programmen begann, wird wohl auch in Zukunft weitergehen – jedoch auf einer völlig neuen Ebene. Die Schweiz nimmt das Thema auf jeden Fall ernst: Im August dieses Jahres ermächtigte der Bundesrat das Departement für Verteidigung, Bevölkerungsschutz und Sport (VBS), einen Antrag auf Mitgliedschaft der European Cyber Security Organisation (ECSO) zu stellen. ECSO ist die einzige europäische sektorübergreifende und unabhängige Mitgliederorganisation für Cybersicherheit. Zudem genehmigte der Bundesrat die Teilnahme des VBS am Projekt «Cyber Ranges Federation» der Permanent Structured Cooperation (PESCO) der EU. Im Rahmen der Strategie Digitale Schweiz beschäftigt man sich mit der Ausarbeitung eines schweizerischen Ansatzes für die Regulierung von KI. Ob sich Angreifer von solchen Regulierungen aufhalten lassen, steht auf einem anderen Bildschirm. ■