KI und Cybersicherheit

Martina Dalla Vecchia und Rainer Kessler

Kapitelzusammenfassung

Führungskräfte sind heute auch **Digital Leaders** – mit einer klaren Verantwortung für den sicheren und strategischen Einsatz von Technologie. Besonders im Bereich **künstliche Intelligenz** (**KI**) und hinsichtlich **Cyber-Bedrohungen** ist Umsicht gefragt. Eine KMU-Führungskraft muss verstehen, dass künstliche Intelligenz die Bedrohungslage für Unternehmen grundsätzlich verändert – zum Guten wie zum Schlechten. KI kann Cyberangriffe effektiver machen, aber auch den Schutz von Unternehmenswerten verbessern. Es geht nicht nur darum, IT-Sicherheitsteams zu unterstützen, sondern auch um strategische Entscheidungen, die Compliance, Risikomanagement und Unternehmensprozesse betreffen. Dieses Kapitel zeigt, wie KI und Cybersicherheit zusammenhängen. Es liefert **klare Handlungsempfehlungen und eine Checkliste für KMU**, um KI sicher und verantwortungsvoll einzusetzen – als Schutzschild statt als Risiko.

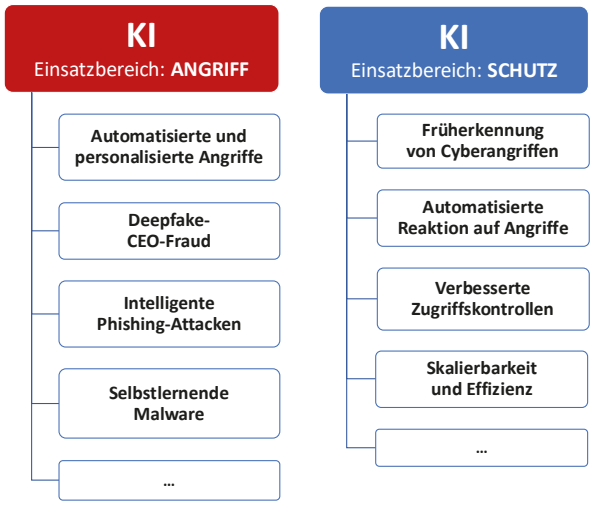

Abbildung 1: Gegenüberstellung der Einsatzbereiche der KI für Angriff und Schutz

Begriffsabgrenzungen

Vorliegend wird der Themenbereich «KI und Cybersicherheit» betrachtet.[1]

Künstliche Intelligenz (KI) wird hierfür auf zwei unterschiedliche Arten definiert. Einerseits umfasst KI technische Systeme, welche mittels maschinellen Lernens eine nicht-determinierte Funktionsweise erlangen und damit Verarbeitungen ermöglichen, welche «intelligent» (im Sinne von von Menschen wahrgenommener Intelligenz, unabhängig von deren Stärke) wirken – das ist eine begrenzte Definition, welche die Verwendung des Begriffs «KI» nur im Zusammenhang mit maschinellem Lernen zulässt. Andererseits wird vielerorts aktuell eine Definition verwendet, welche alle Systeme einbezieht, welche «intelligent» (im Sinne von von Menschen wahrgenommener Intelligenz, unabhängig von deren Stärke) wirken – das ist eine erweiterte Definition, welche u. a. auch Expertensysteme mit determiniertem Funktionsumfang umfasst, sofern diese im vorgenannten Sinne «intelligent» wirken. Es sind noch weitere Definitionen denkbar und auch im Umlauf, doch vorliegend beschriebene zwei Definitionen sind in der Praxis häufig anzutreffen und ermöglichen den effizienten Umgang mit dem Thema.[2] Im vorliegenden Kapitel wird mit der «begrenzten Definition» gearbeitet: Wir nennen also etwas nur dann KI, wenn die zu Grunde liegende Technologie das maschinelle Lernen ist.

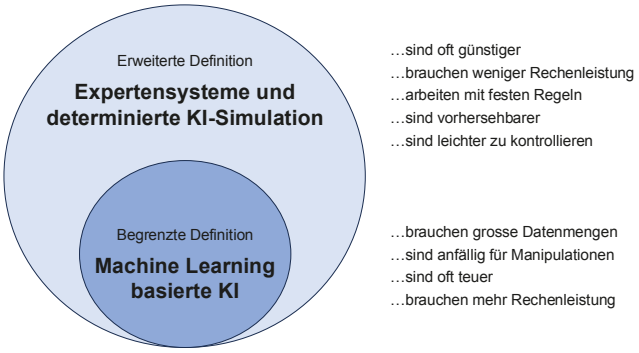

Abbildung 2: Begriffsabgrenzung künstliche Intelligenz

1 Die im gesamten Kapitel genannten Quellen sind generell als weiterführend zu verstehen und enthalten nicht zwingend die verwendeten Informationen.

2 Bitkom e.V. und Deutsches Forschungszentrum für künstliche Intelligenz.

Es ist wichtig, genau zu definieren, um welche Art von KI es sich handelt, da sonst ein erhöhtes Risiko eines Verstosses gegen Artikel 22 DSGVO oder Artikel 21 CH-DSG besteht. Und es kann in der Tat zu falschen Erwartungen oder sogar Fehlentscheiden führen, da je nach Technologie die Antwortverlässlichkeit unterschiedlich ist.

Cybersicherheit ist der zweite relevante Begriff für dieses Kapitel. Wir verstehen darunter die Sicherheit von Informatik- und Telematik-Systemen mit Schwerpunkt Resistenz gegen absichtliche Angriffe oder Kollateralschäden von solchen. In der Praxis gibt es – wie für den KI-Begriff – viele gebräuchliche Definitionen. Entsprechend gilt es hier etwas flexibel zu sein in der Interpretation. Je nach Anwendungsfall können mit Cybersicherheit auch Sicherheitsmassnahmen umschrieben werden, welche nicht nur für die Erkennung, Abwehr und Nachbearbeitung von Angriffen geeignet sind, sondern die generelle Sicherheit der IKT ermöglichen (z. B. auch gegen Stromausfall oder Erdbeben). Eine extrem breite Definition des Cybersicherheitsbegriffs ist im Kontext kritischer Infrastruktur (Wasser- und Nahrungsversorgung, Energie, Gesundheitseinrichtungen, Blaulichtorganisationen, Entsorgung, Informationsverbreitung, Finanzwesen, etc.) gebräuchlich: Es handelt sich dabei um die Sicherheit von sogenannten «Cyber-Physical-Systems» – gemeint ist damit nicht nur die Sicherheit im Cyberspace, sondern auch die Sicherheit der Anlagen selbst.[3] Diese Definition ist jedoch für KMU selten relevant, weshalb wir uns auf die Sicherheit der IKT und die Abwehr von Angriffen konzentrieren.

KI als Sicherheitsrisiko – Cyberangriffe werden «intelligenter»

Mit der fortschreitenden Entwicklung der künstlichen Intelligenz (KI) eröffnen sich nicht nur neue Möglichkeiten für Innovation und Wachstum, sondern auch bedrohliche Perspektiven für die Cybersicherheit. Hacker:innen und Cyberkriminelle nutzen zunehmend KI-Technologien, um ihre Angriffe zu optimieren und zu automatisieren. Diese Entwicklung hat tiefgreifende Auswirkungen auf die Art und Weise, wie Unternehmen ihre digitalen Vermögenswerte schützen müssen.

Hacker:innen setzen KI gezielt ein: KI ermöglicht automatisierte und personalisierte Angriffe – von **Deepfake-CEO-Fraud** über **intelligente Phishing-Attacken** bis hin zu **selbstlernender Malware**. Die Fähigkeit der KI, grosse Datenmengen in kürzester Zeit zu analysieren, bietet Angreifer:innen die Möglichkeit, Schwachstellen in IT-Systemen schneller und effektiver zu identifizieren und auszunutzen. Zudem können personalisierte Angriffsmethoden entwickelt werden, die auf das Verhalten und die Vorlieben einzelner Nutzer:innen zugeschnitten sind, wodurch die Erfolgsquote solcher Angriffe steigt.[4]

Erhöhte Geschwindigkeit und Skalierbarkeit: KI kann grosse Datenmengen analysieren und gezielt Sicherheitslücken ausnutzen. Ein bedeutender Vorteil der KI ist ihre Fähigkeit zur Skalierung. Während ein:e menschliche:r Hacker:in nur eine begrenzte Anzahl von Angriffen gleichzeitig ausführen kann, ermöglicht die Automatisierung durch KI die Durchführung einer Vielzahl von Angriffen parallel. Dies erhöht nicht nur die Reichweite, sondern auch die Geschwindigkeit, mit der Angriffe stattfinden können. Ein bekanntes Beispiel ist die Verwendung von KI-gestützter Malware, die sich selbst weiterentwickelt und anpasst, um den Verteidigungsmechanismen zu entgehen. Zudem kann sich solche Malware «tarnen», sodass die Schaden-Funktion nicht sichtbar ist, bis ein bestimmter System-Zustand (z. B. durch KI identifiziertes User-Verhalten, bis hin zu mit KI erstellten psychologischen oder politischen Profilen) erreicht wird. Erst dann entschlüsselt sich die Schaden-Funktion automatisch und entfaltet ihre böswillige Wirkung (ein Beispiel für eine solche weit entwickelte Malware ist «Deeplocker»).[5]

3 IT-Sicherheit; Konzepte – Verfahren – Protokolle; 11. Auflage; Prof. Dr. Claudia Eckert; De Gruyter, Oldenburg, 2023.

4 BSI – KI und gegenwärtige Cyberbedrohungen (www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Cyber-Sicherheitslage/Analysen-und-Prognosen/Threat-Intelligence/Blogeintraege/KI_gegenwaertige-Cyberbedrohungen.html).

5 ZDNET (www.zdnet.de/88339715/deeplocker-ibm-entwickelt-auf-kuenstlicher-intelligenz-basierende-malware).

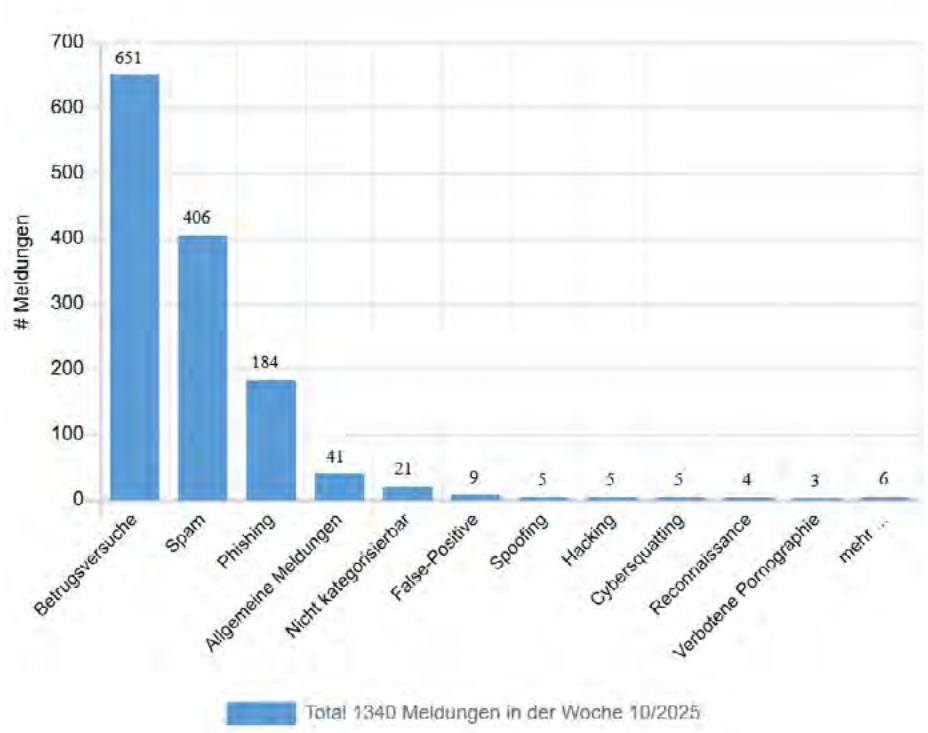

Abbildung 3: Cyber-Angriffe in der Schweiz: Meldeeingang Woche 10/2025 (www.ncsc.ch)

Vertrauensprobleme durch **Fake-Content**: KI-generierte Deepfakes und gefälschte Texte untergraben die Vertrauenswürdigkeit von Kommunikation und Daten. Deepfake-Technologien ermöglichen es, täuschend echte Videos und Audiodateien zu erstellen, die kaum von echten Aufnahmen zu unterscheiden sind. Diese gefälschten Inhalte können dazu verwendet werden, Desinformationen zu verbreiten oder Personen zu erpressen. Besonders gefährlich ist dies, wenn hochrangige Unternehmensvertreter:innen oder politische Persönlichkeiten ins Visier genommen werden. Auch die Verbreitung von gefälschten E-Mails und Dokumenten kann zu erheblichen Schäden führen, indem sie das Vertrauen in digitale Kommunikation untergräbt.[6]

Ein Beispiel für einen erfolgreichen KI-gestützten Angriff ist die Verwendung von KI zur Durchführung von **Spear-Phishing-Angriffen**. Hierbei analysiert die KI das Verhalten und die Kommunikationsmuster eines Unternehmens, um massgeschneiderte Phishing-E-Mails zu erstellen, die von den Empfänger:innen als legitim wahrgenommen werden. Ein weiteres Beispiel ist die Nutzung von KI in Botnetzen, die dazu verwendet werden, koordinierte Angriffe auf Netzwerke und Websites durchzuführen, wodurch diese überlastet und funktionsunfähig gemacht werden.[7]

Angesichts dieser Bedrohungen müssen Unternehmen ihre Abwehrstrategien kontinuierlich weiterentwickeln und anpassen. Dabei spielt die Integration von KI in Sicherheitslösungen eine entscheidende Rolle. KI kann dabei helfen, Anomalien im Netzwerkverkehr zu erkennen, verdächtige Aktivitäten zu identifizieren und automatisch Gegenmassnahmen einzuleiten. Dies erfordert jedoch eine enge Zusammenarbeit zwischen IT-Sicherheitsteams und der Unternehmensführung, um sicherzustellen, dass die eingesetzten Technologien auf dem neuesten Stand sind und effektiv genutzt werden.

6 UZH (www.foeg.uzh.ch/de/research/projects/deep-fakes-wahrnehmung.html).
7 Täglich updatete Angriffsbeschreibungen: www.databreachtoday.com (englisch).

Insgesamt zeigt sich, dass künstliche Intelligenz sowohl eine grosse Chance als auch ein erhebliches Risiko für die Cybersicherheit darstellt. Unternehmen müssen sich der Gefahren bewusst sein und proaktive Massnahmen ergreifen, um sich gegen die zunehmende Bedrohung durch KI-gestützte Cyberangriffe zu wappnen.

KI als Sicherheitslösung – Schutz durch intelligente Systeme

Künstliche Intelligenz (KI) bietet nicht nur Angriffsmöglichkeiten, sondern auch erhebliche Potenziale, um die Cybersicherheit zu verbessern. Wenn Unternehmen KI in ihre Sicherheitsstrategien integrieren, verbessern sich ihre Chancen, Bedrohungen entgegenzuwirken und Angriffe abzuwehren.

Früherkennung von Cyberangriffen: Ein entscheidender Vorteil von KI ist ihre Fähigkeit zur Anomalie-Erkennung. Durch die kontinuierliche Überwachung des Netzwerkverkehrs und die Analyse von Datenmustern sowie des Verhaltens von User:innen kann KI ungewöhnliches Verhalten frühzeitig identifizieren. Dabei geht es nicht in erster Linie um Vertrauen in die eigenen Mitarbeiter:innen, sondern darum, zu erkennen, wenn deren digitale Identität von Angreifer:innen übernommen wurde. Diese Predictive Security erlaubt es Unternehmen, potenzielle Bedrohungen zu erkennen, bevor sie Schaden anrichten können.[8]

Automatisierte Reaktion auf Angriffe: KI-gestützte Sicherheitssysteme sind in der Lage, Sicherheitsvorfälle in Echtzeit zu analysieren und automatisch Gegenmassnahmen einzuleiten. Dies erfolgt durch die Nutzung von Machine Learning-Algorithmen, die aus vergangenen Angriffen lernen und sich kontinuierlich verbessern. Sobald eine Bedrohung erkannt wird, können diese Systeme automatisch Schutzmassnahmen wie das Isolieren betroffener Bereiche oder das Blockieren verdächtiger Aktivitäten durchführen.[9]

Verbesserte Zugriffskontrollen: Durch den Einsatz von KI können Unternehmen ihre Zugriffskontrollen deutlich verbessern. Technologien wie Gesichtserkennung und Verhaltensanalyse ermöglichen eine präzise Authentifizierung der Benutzer:innen. Dies erhöht die Sicherheit, indem unerlaubte Zugriffe frühzeitig erkannt und verhindert werden. KI-gestützte Authentifizierungstechniken bieten eine zusätzliche Sicherheitsschicht, die traditionelle Passwortsysteme ergänzt.[10]

Skalierbarkeit und Effizienz: Ein weiterer Vorteil von KI in der Cybersicherheit ist ihre Fähigkeit zur Skalierung. Während menschliche Expert:innen nur eine begrenzte Anzahl von Bedrohungen gleichzeitig analysieren können, ermöglicht die Automatisierung durch KI die gleichzeitige Überwachung und Analyse einer Vielzahl von Datenquellen. Dies erhöht die Effizienz und ermöglicht es Unternehmen, auch bei begrenztem Budget effektive Sicherheitsmassnahmen zu implementieren.[11]

KI-basierte Sicherheitssysteme sind nicht nur für grosse Unternehmen von Vorteil. Auch kleine und mittelständische Unternehmen (KMU) können von den fortschrittlichen Sicherheitslösungen profitieren, die durch KI ermöglicht werden. Durch den Einsatz von Cloud-basierten Sicherheitslösungen können auch KMU ihre Netzwerke und Daten effektiv schützen, ohne hohe Investitionen in Hardware und Personal tätigen zu müssen.[12]

Insgesamt stellt die **Integration von KI in Sicherheitsstrategien** eine vielversprechende Möglichkeit dar, um den Herausforderungen der modernen Cybersicherheitslandschaft zu begegnen. Unternehmen müssen jedoch sicherstellen, dass ihre KI-gestützten Sicherheitslösungen auf dem neuesten Stand sind und kontinuierlich weiterentwickelt werden, um den sich ständig ändernden Bedrohungen gerecht zu werden. Da dies für ein KMU eine Herausforderung darstellt, kann – wie im vorhergehenden Absatz erläutert – auf Cloud-basierte Sicherheitslösungen zurückgegriffen werden. Im nächsten Kapitel gehen wir noch detaillierter darauf ein, wie KI konkret eingesetzt werden kann, um Risiken zu identifizieren und diesen zu begegnen.

8 Predictive Security: Frühzeitige Erkennung und Abwehr durch KI.
9 Machine Learning: Algorithmen, die aus Daten lernen und sich kontinuierlich verbessern.
10 KI-gestützte Authentifizierungstechniken: Technologien zur präzisen Benutzererkennung.
11 Skalierbarkeit: Fähigkeit von KI, grosse Datenmengen effizient zu verarbeiten.
12 Cloud-basierte Sicherheitslösungen: KI-Sicherheitslösungen auch für KMU.

Risikomanagement – Gefahren und Chancen der eigenen KI-Anwendung

End-User-KI-Anwendungen (generative KI)

Der Einsatz von generativer KI durch Mitarbeiter:innen eines Unternehmens bietet sowohl Chancen als auch Risiken, die sorgfältig abgewogen werden müssen. Einer der grössten Vorteile besteht in der erheblichen Steigerung der Produktivität und Kreativität der Mitarbeiter:innen. Generative KI kann dabei unterstützen, Routineaufgaben zu automatisieren, komplexe Datenanalysen durchzuführen und neue, innovative Ideen zu generieren. Dies ermöglicht es den Mitarbeiter:innen, sich auf anspruchsvollere und wertschöpfende Tätigkeiten zu konzentrieren. Zudem können personalisierte Inhalte und Lösungen in Echtzeit erstellt werden, was die Effizienz und Kundenzufriedenheit erhöht.[13]

Dennoch bringt die Nutzung von generativer KI auch Herausforderungen und potenzielle Gefahren mit sich. Eine der grössten Gefahren besteht in der Möglichkeit von Fehlinformationen oder ungenauen Ergebnissen. Wenn generative KI auf unzureichenden oder verzerrten Daten basiert, können die resultierenden Vorhersagen und Inhalte irreführend sein. Dies könnte zu falschen Entscheidungen und erheblichen geschäftlichen Risiken führen. Darüber hinaus besteht die Gefahr, dass sensible Unternehmensdaten missbraucht oder durch die KI ungewollt offengelegt werden.

Ein weiteres Risiko ist der mögliche Missbrauch von generativer KI durch Mitarbeiter:innen. Ohne angemessene Richtlinien und Überwachungsmechanismen könnten KI-Tools für unethische oder illegale Aktivitäten genutzt werden, was erhebliche rechtliche und imagebezogene Konsequenzen nach sich ziehen könnte. Aus diesem Grund ist es entscheidend, dass Unternehmen klare Nutzungsrichtlinien für generative KI entwickeln und ihre Mitarbeiter:innen entsprechend schulen. Solche Richtlinien können bspw. verbieten, Personendaten oder vertrauliche Informationen mit GenAI-Tools zu bearbeiten.

Zusammenfassend lässt sich sagen, dass die Chancen, die durch den Einsatz von generativer KI geboten werden, erheblich sind, allerdings nur dann voll ausgeschöpft werden können, wenn gleichzeitig auch die damit verbundenen Risiken angemessen adressiert werden. Unternehmen sollten einen proaktiven Ansatz verfolgen, um sicherzustellen, dass die Vorteile der generativen KI maximiert und die potenziellen Gefahren minimiert werden. Dies erfordert eine kontinuierliche Überwachung, Schulung und Anpassung der KI-Strategien an die sich wandelnden Gegebenheiten und Bedrohungen.

Anwendungen des maschinellen Lernens

Maschinelles Lernen bietet zwar vielfältige Anwendungsmöglichkeiten und ist aus dem Arbeitsalltag vielerorts nicht mehr wegzudenken, jedoch sind die damit verbundenen Risiken und Gefahren nicht zu unterschätzen. Unternehmen müssen sicherstellen, dass ihre eigenen KI-Modelle nicht manipuliert oder getäuscht werden, beispielsweise durch **Adversarial Attacks** oder **Datenvergiftung**, die die Ergebnisse verfälschen und erhebliche Schäden verursachen können. Die Datenqualität ist hierbei entscheidend, denn KI-Systeme sind nur so gut wie ihre Trainingsdaten. Wenn diese verfälscht oder nicht repräsentativ sind, entstehen Fehleinschätzungen, die zu falschen Entscheidungen und finanziellen Verlusten führen können.[14]

Ein weiteres Risiko besteht darin, dass automatisierte oder durch maschinelles Lernen unterstützte Entscheidungen von Interessengruppen als **diskriminierend** wahrgenommen werden können, selbst wenn diese auf statistischen Analysen basieren. Dies kann zu rechtlichen und reputationsbezogenen Konsequenzen führen. Daher ist es wichtig, vor der Nutzung von KI-Output eine kritische und schriftlich dokumentierte Beurteilung der möglichen Folgen vorzunehmen, um potenzielle Risiken zu minimieren.

Unternehmen, die **KI-Systeme von Drittanbietern** nutzen, sollten sicherstellen, dass diese transparent, vertrauenswürdig und robust (Verfügbarkeit) sind. Die Auswahl eines geeigneten Anbieters und die regelmässige Überprüfung der Systeme sind unerlässlich, um die Integrität und Verlässlichkeit der KI-Anwendungen zu gewährleisten.

Ein weiteres Problem ist die sogenannte **«Black-Box»-Natur** vieler KI-Modelle, bei der die Entscheidungsfindung für menschliche Benutzer:innen nicht immer nachvollziehbar ist. Dies kann Misstrauen und Bedenken hinsichtlich der Fairness und Genauigkeit der Entscheidungen wecken. Transparenz und Erklärbarkeit sind daher wichtige Aspekte,

13 Digital.Now (www.valantic.com/de/digital-now/der-grosse-genai-report-status-quo-risiken-und-visionen-deutscher-unternehmen).

14 T-Systems (www.t-systems.com/ch/de/whitepaper-download/qualitaetssicherung-entwicklung-ki-systeme-413002).

die bei der Entwicklung und Implementierung von KI-Systemen berücksichtigt werden müssen. Unternehmen sollten darauf achten, dass ihre Modelle nicht nur leistungsfähig, sondern auch erklärbar sind, um Vertrauen bei den Nutzer:innen zu schaffen und die Akzeptanz der Technologie zu fördern.

Ein spezifisches Anwendungsgebiet des maschinellen Lernens, das häufig übersehen wird, ist die **vorausschauende Wartung**. Anhand von Sensordaten und historischen Wartungsaufzeichnungen können Maschinen lernen, wann ein Gerät wahrscheinlich ausfallen wird. Dies ermöglicht es Unternehmen, proaktiv Wartungsarbeiten durchzuführen und Ausfallzeiten zu minimieren. Dadurch können die Effizienz gesteigert und die Betriebskosten gesenkt werden. Allerdings müssen Unternehmen sicherstellen, dass die dabei verwendeten Daten akkurat und umfassend sind, um verlässliche Vorhersagen zu ermöglichen.

Zudem kann die **Bilderkennung** Chancen und Risiken bergen. Maschinelles Lernen kann eingesetzt werden, um Bilder und Videos zu analysieren und Objekte, Personen oder Szenen zu identifizieren. Dies ist besonders nützlich in der Sicherheitsbranche, wo Überwachungskameras mit maschinellem Lernen ausgestattet werden können, um Sicherheitsbedrohungen in Echtzeit zu erkennen. Auch in der Gesundheitsbranche findet diese Technologie Anwendung, beispielsweise zur Analyse von medizinischen Bildern zur frühzeitigen Erkennung von Krankheiten. Die Genauigkeit und Zuverlässigkeit der Bildanalysen hängen jedoch stark von der Qualität und Repräsentativität der Trainingsdaten ab, was erneut die Bedeutung der Datenqualität unterstreicht.[15]

Zusammenfassend lässt sich sagen, dass maschinelles Lernen zwar erhebliche Vorteile bietet, jedoch ein sorgfältiger und verantwortungsbewusster Umgang mit dieser Technologie erforderlich ist, um die damit verbundenen Risiken zu minimieren. Unternehmen sollten kontinuierlich in die **Qualität ihrer Daten** und die **Integrität ihrer KI-Systeme** investieren, um die bestmöglichen Ergebnisse zu erzielen und gleichzeitig die Gefahren zu beherrschen. Dies erfordert eine umfassende Strategie, die sowohl technische als auch ethische Aspekte berücksichtigt und sicherstellt, dass die eingesetzten KI-Modelle transparent, korrekt, zuverlässig und fair sind.

Eigenes Risikoprofil und passende Massnahmen mit KI erarbeiten

Allgemeines
Die Nutzung von künstlicher Intelligenz (KI) bietet Unternehmen die Möglichkeit, ein umfassendes und differenziertes Risikoprofil zu erstellen. Dies ist nicht nur ein wertvolles Werkzeug für grosse Konzerne, sondern kann auch kleinen und mittelständischen Unternehmen (KMU) dabei helfen, ihre Sicherheitslücken zu identifizieren und gezielte Massnahmen zur Risikominimierung zu entwickeln.[16]

Ermittlung von Unternehmenswerten
Durch den Zugriff auf Open-Source-Informationen und interne Daten kann die KI wichtige Unternehmenswerte identifizieren, die zum Erfolg der Organisation beitragen. Diese Analyse ermöglicht es, strategische Prioritäten zu setzen und den Schutz der wertvollsten Ressourcen sicherzustellen.

Identifikation von Schwachstellen
Weiter kann die KI Schwachstellen und potenzielle Angriffspunkte aufdecken. Dies ist besonders wichtig, um gezielte Sicherheitsmassnahmen zu ergreifen und das Unternehmen vor Cyberangriffen zu schützen. Hierbei können aktuelle Bedrohungen aus dem Internet in Echtzeit berücksichtigt werden, was eine dynamische und aktuelle Schwachstellenidentifikation ermöglicht.

Erstellung eines Bedrohungsbildes
Basierend auf den gesammelten Daten kann die KI ein detailliertes Bedrohungsbild erstellen. Sie analysiert aktuelle Cyberangriffe, die in den Medien dokumentiert sind, und vergleicht diese mit den internen Schwachstellen des Unternehmens – wo akute Bedrohungen auf Schwachstellen treffen, besteht eine besonders grosse Gefahr. Dies führt zu einem spezifischen und relevanten Risikoinventar.

Risikoanalyse und Massnahmenvorschläge
Die KI bewertet die einzelnen Risiken und erstellt eine Prioritätenliste für Massnahmen zur Risikominderung. Diese Vorschläge können den spezifischen Bedürfnissen und Anforderungen des Unternehmens angepasst werden. Zudem kann die KI Steuerungsparameter wie z. B. Risikoappetit und Risikotoleranz berücksichtigen, um eine individuelle Risiko-Priorisierung zu gewährleisten.

15 ARGUS Data Insights (www.argusdatainsights.ch/de/blog/die-macht-der-bilderkennung).
16 FHNW CAS AI powered CyberTech (www.fhnw.ch/de/weiterbildung/wirtschaft/cas-ai-powered-cybertech).

Kritische Überprüfung und Dokumentation
Die Ergebnisse der KI-Analyse sollten stets kritisch hinterfragt und dokumentiert werden. Die Verantwortung für das Risikomanagement liegt schlussendlich beim Management des Unternehmens. Eine sorgfältige Überprüfung und Anpassung der Vorschläge der KI ist daher unerlässlich, um sicherzustellen, dass sie den tatsächlichen Bedürfnissen und Einschätzungen der Unternehmensführung entsprechen.

Dieses differenzierte Vorgehen ermöglicht es Unternehmen, ein robustes Risikomanagement zu etablieren, das sowohl technische als auch ethische Aspekte berücksichtigt und die Sicherheit und Resilienz der Organisation stärkt – und zwar fast gänzlich automatisch mit KI, was das Vorgehen gerade für KMU attraktiv macht.

Schulung der Mitarbeiter:innen

Sicherer Umgang mit KI-Tools

Die Schulung der Mitarbeiter:innen im sicheren Umgang mit KI-Tools ist von zentraler Bedeutung, um die Effektivität und Sicherheit der KI-Anwendungen im Unternehmen zu gewährleisten. Zu Beginn der Schulung sollte den Mitarbeiter:innen ein umfassendes Verständnis der grundlegenden Funktionsweisen von KI-Tools vermittelt werden. Dies schliesst eine Einführung in die verschiedenen Arten von KI, deren Einsatzmöglichkeiten und deren Grenzen ein.

Ein weiterer wichtiger Bestandteil der Schulung ist die **Sensibilisierung** für mögliche Risiken und Schwachstellen, die mit dem Einsatz von KI-Tools einhergehen können. Dabei sollten konkrete Fallbeispiele und Szenarien durchgesprochen werden, um die Mitarbeiter:innen auf potenzielle Bedrohungen wie Datenmissbrauch, Fehlentscheidungen durch fehlerhafte Algorithmen oder unerwünschte Verhaltensweisen von KI-Systemen vorzubereiten.

Zur Vertiefung dieses Wissens ist es wichtig, den Mitarbeiter:innen spezifische Sicherheitsmassnahmen an die Hand zu geben. Dazu zählen **Anweisungen** zur sicheren Konfiguration und Nutzung der KI-Tools, regelmässige Updates und Patches sowie der Umgang mit sicherheitsrelevanten Vorfällen. Auch die Bedeutung eines sicheren Passwortmanagements sowie der Schutz sensibler Daten sollten hervorgehoben werden.

Abschliessend ist es essenziell, dass die Mitarbeiter:innen regelmässig über die aktuellen Richtlinien und Vorschriften im Umgang mit KI-Tools informiert werden. Dies umfasst auch die Einhaltung der gesetzlichen Bestimmungen und internen Weisungen sowie die Berücksichtigung ethischer Aspekte bei der Nutzung von KI-Systemen.

Cybersicherheitsschulung mit KI entwickeln

Die **Schulung der Mitarbeiter:innen** im Bereich Cybersicherheit ist ein wesentlicher Bestandteil des **Risikomanagements** und unterscheidet sich thematisch vom sicheren Umgang mit KI-Tools. Während der Fokus des vorhergehenden Kapitels auf der Nutzung und Sicherheit von KI-Anwendungen liegt, befasst sich die Cybersicherheitsschulung umfassender mit der Abwehr von Bedrohungen und dem Schutz der IT-Infrastruktur.

Der erste Schritt beim Aufbau einer Cybersicherheitsschulung ist die Erstellung eines umfassenden Schulungsplans, der alle relevanten Themen abdeckt. Dazu gehören die Grundlagen der Cybersicherheit wie zum Beispiel die Erkennung und Abwehr von Phishing-Angriffen, die Bedeutung sicherer Passwörter, der Schutz sensibler Daten und die sichere Nutzung von Netzwerken und Geräten. Ziel ist es, den Mitarbeiter:innen ein ganzheitliches Verständnis für die Cyberbedrohungen zu vermitteln, denen sie täglich ausgesetzt sein könnten.

Ein zentraler Bestandteil der Schulung ist das Training spezifischer **Sicherheitsmassnahmen**. Dies kann beispielsweise die regelmässige Aktualisierung von Software und Systemen, das Erkennen und Melden verdächtiger Aktivitäten sowie die korrekte Handhabung von Sicherheitsvorfällen umfassen. Durch praktische Übungen und Simulationen sollen die Teilnehmer:innen in die Lage versetzt werden, realitätsnahe Bedrohungsszenarien zu erkennen und angemessen darauf zu reagieren.

Die Schulungen sollte auch interaktive Elemente enthalten, um das Engagement der Mitarbeitenden zu fördern und das Gelernte zu festigen. Rollenspiele, Workshops und Fallstudien können helfen, theoretisches Wissen in die Praxis umzusetzen. Zudem ist es wichtig, die Schulungen regelmässig zu wiederholen und an aktuelle Bedrohungen und technologische Entwicklungen anzupassen.

Abschliessend ist es von grosser Bedeutung, dass die Mitarbeiter:innen die geltenden gesetzlichen Vorschriften und internen Richtlinien im Bereich Cybersicherheit kennen und einhalten. Dies umfasst beispielsweise die Datenschutzgesetze und die internen Weisungen zum Umgang mit vertraulichen Informationen. Eine **kontinuierliche Sensibilisierung** und Schulung der Mitarbeiter:innen trägt dazu bei, das Sicherheitsbewusstsein zu stärken und die **Resilienz** des Unternehmens gegen Cyberangriffe zu erhöhen.[17]

Rechtliche Anforderungen und Compliance

KI, Cybersicherheit und Recht allgemein

Für KI und Cybersicherheit gibt es National und international diverse Gesetze, Verordnungen und Regulatorien sowie Standards, welche aufgrund ihrer grossen Verbreitung eine defacto bindende oder zumindest wegweisende Rolle spielen.

Führungskräfte, welche zulassen, dass in ihrem Betrieb solche «Wegweiser» missachtet werden, können ggf. zur Verantwortung gezogen werden. Gleichzeitig kann man von einer Führungskraft nicht verlangen, dass sie oder er alle einschlägigen Gesetze in diesem Bereich kennt. Wie also umgehen mit diesem Dilemma? Der gesunde Menschenverstand und eine solide Ethik sind **gute Mittel, um mindestens 80 % des Weges zur Gesetzeseinhaltung** zu meistern. Doch kann es sich ein KMU leisten, die restlichen 20 % dem Zufall zu überlassen? Wohl eher nicht – zumindest nicht als offizielle Politik im Umgang mit den Themen KI und Cybersicherheit.[18]

Die nachfolgenden Informationen dieses Kapitels helfen, den relevanten Teil der letzten 20 % zur vollständigen Compliance zu erschliessen.

Datenschutz

Über Datenschutz und KI sowie über Datenschutz und Cybersicherheit und sogar über Datenschutz und die Kombination aus KI und Cybersicherheit kann man viel schreiben, denn es gibt viele meist indirekte Beeinflussungsmöglichkeiten von Personendaten durch diese Themenbereiche. Interessant wird es, wenn die Thematik auf das absolute Minimum reduziert wird: Wo trifft der Datenschutz genau und unmittelbar auf die KI und auf die Cybersicherheit – und umgekehrt? Die Antwort ist: z. B. in den Artikeln 22 und 32 der Datenschutzgrundverordnung der EU (DSGVO) sowie in den Artikeln 21 und 8 des Schweizer Datenschutzgesetzes. Dabei entsprechen sich die Artikel 22 DSGVO und 21 DSG weitgehend, denn das (neue) Schweizer Datenschutzgesetz ist eine Reaktion auf die DSGVO und es ist Voraussetzung dafür, dass ein recht freier Datenaustausch zwischen der Schweiz und der EU stattfinden kann. Gleiches gilt für die Artikel 32 DSGVO und 8 DSG. Die Artikel 22, resp. 21 beziehen sich unmittelbar auf KI und die Artikel 32 resp. 8 beziehen sich unmittelbar auf die Cybersicherheit.

Der Datenschutz verlangt beim Einsatz von KI Transparenz gegenüber den Personen, über die Personendaten mit KI verarbeitet werden – und zwar dann, wenn für diese Personen ein relevantes Risiko mit der Verarbeitung zusammenhängt (z. B. automatische Bonitätsprüfung oder Kundenprofilbildung mit nachfolgender Einschränkung der Service-Möglichkeiten). Der Datenschutz nennt dies eine automatisierte Entscheidung im Einzelfall. Unternehmen – auch KMU – müssen also wissen, wo bei ihnen KI im Zusammenhang mit Personendaten zum Einsatz kommt. Diese personendatenbezogenen KI-Anwendungen müssen nicht nur inventarisiert, sondern hinsichtlich ihres Risikos für die betroffenen Personen gewichtet sein. Ist das Risiko relevant, gilt es die betroffenen Personen in geeigneter Form zu informieren und ihnen die Möglichkeit zu geben, aktiv zu werden und eine alternative Verarbeitung (ohne KI) zu verlangen. Gerade in digitalisierten Geschäftsmodellen kann das bspw. aufgrund der ggf. grossen Anzahl Kund:innen eine Herausforderung sein. Hier muss praktisch gedacht werden: Der Interventionsprozess muss ebenso digitalisiert werden. Die «nicht im Einzelfall automatisierte» Alternative eines Prozesses kann jedoch nach wie vor softwaregestützt ablaufen, z. B. indem eine weniger riskante Entscheidung ohne maschinelles Lernen, sondern durch einen nachweislich von Menschen entwickelten determinierten Entscheidungsbaum angewandt wird. Wenn das nicht geht, dann muss im Einzelfall ein Mensch die Beurteilung vornehmen.

17 FHNW CAS Cybersicherheit und Information Risk Management (CSIRM, www.fhnw.ch/de/weiterbildung/wirtschaft/cas-cybersecurity-und-information-risk-management).

18 iusnet – digitales Recht und Datenrecht (digitalesrecht-datenrecht.iusnet.ch/de/stichwortverzeichnis/cybersecurity).

Im Bereich der Cybersicherheit verlangt der Datenschutz, dass angemessene (betr. ihrer Gestaltung) und wirksame (betr. ihres Betriebs) Sicherheitsmassnahmen zum Schutz von Personendaten umgesetzt und wirksam sind. Das Gesetzt spricht hier von technischen und organisatorischen Massnahmen (TOM). Auch hier gibt es gute Vorlagen, die sogar mehr als die oben erwähnten 80 % (mit gesundem/normalem Menschenverstand machbar) erschliessen. Es handelt sich dabei um Standards, wie bspw. ISO 27001[19] und ISO 27002[20] sowie NIST CSF 2.0[21], aber auch um viele weitere z. T. branchenspezifische Standards. Wichtig ist es, dass der Inhalt von solchen Standards nicht ohne Reflexion umgesetzt wird, sondern die Liste der meist mehrere Duzend Massnahmenvorschläge hinterfragt wird und in der Folge eine Begründung für diejenigen Massnahmen dokumentiert wird, welche im eigenen Fall nicht anwendbar sind. Zudem ist die Gesamtheit der TOM in einer gewissen Regelmässigkeit hinsichtlich ihrer Wirksamkeit und Aktualität sowie Angemessenheit zu hinterfragen – hierüber ist Buch zu führen. Wenn diese Vorgänge nachvollziehbar durchgeführt wurden, so vereinfacht sich die Argumentation ungemein, dass ein allfälliger Cyber-Vorfall auf ein «**richtigerweise akzeptiertes Restrisiko**» zurückzuführen ist und entsprechend **keine Haftung für die Führungs-Crew** ansetzbar ist.

Weisungswesen für KI

Um den Einsatz von KI in Unternehmen effektiv zu steuern, ist ein umfassendes **Weisungswesen** unerlässlich. Dieses sollte klare Richtlinien für die Implementierung und Nutzung von KI-Technologien beinhalten. Dazu gehört die Festlegung von Verantwortlichkeiten, die Definition von Prozessen zur Überwachung und Bewertung der KI-Anwendungen sowie regelmässige Audits, um die Einhaltung der Vorgaben sicherzustellen. Unternehmen sollten auch sicherstellen, dass Mitarbeiter:innen in diesen Richtlinien geschult werden, um eine korrekte und verantwortungsvolle Nutzung der KI-Systeme zu gewährleisten.

Richtlinien für die Offenlegung der KI-Verwendung

Transparenz ist ein wichtiger Aspekt beim Einsatz von KI. Unternehmen sollten klare **Richtlinien** für die Offenlegung der KI-Verwendung erstellen. Dies umfasst die Information der Nutzer:innen darüber, wann und wie KI-Technologien eingesetzt werden, sowie die Art und Weise, wie Daten gesammelt und verarbeitet werden. Solche Richtlinien fördern das Vertrauen der Nutzer:innen in die KI-Systeme und helfen, ethische Bedenken zu adressieren. Eine transparente Kommunikation trägt dazu bei, Missverständnisse zu vermeiden und die Akzeptanz von KI-Anwendungen zu erhöhen.

Qualitätssicherung von durch KI erstelltem Output

Die Qualitätssicherung des durch KI erstellten Outputs ist von grosser Bedeutung, um die Zuverlässigkeit und Genauigkeit der Ergebnisse zu gewährleisten. Dies erfordert die Implementierung von Mechanismen zur **kontinuierlichen Überprüfung und Validierung** der KI-Modelle. Unternehmen sollten sicherstellen, dass die Algorithmen regelmässig aktualisiert und auf ihre Leistungsfähigkeit hin überprüft werden. Darüber hinaus ist es wichtig, **menschliche Expert:innen** in den Überprüfungsprozess einzubeziehen, um eventuelle Verzerrungen oder Fehler zu identifizieren und zu korrigieren. Eine strenge Qualitätssicherung trägt dazu bei, das **Vertrauen** in die KI-gestützten Systeme zu stärken und ihre Effektivität zu maximieren.

19 Informationssicherheits-Managementstandard der internationalen Standardisierungsorganisation.

20 Liste der einzelnen Sicherheitsmassnahmen, inkl. deren Erläuterung, welche zusammenfassend als Anhang auch im Standard 27001 vorhanden ist.

21 Cybersecurity-Framework des National Institute of Standards and Technology (USA) in der Version 2.0 gilt als einer der wichtigsten Standards für Cybersicherheit weltweit (u. a. orientieren sich der Schutz kritischer Infrastrukturen in der Schweiz sowie die Finanzmarktaufsicht an diesem Rahmenwerk).

Internes Kontrollsystem (IKS)

Relevante ergriffene Massnahmen im Zusammenhang mit der **Sicherheit** oder der **Compliance** in den Themenbereichen KI und Cybersicherheit sollten nicht nur lokal (d. h., in der jeweiligen Prozessdokumentation) dokumentiert sein, sondern auch zentral im Rahmen des **internen Kontrollsystems** (**IKS**) festgeschrieben sein. Das IKS stellt hauptsächlich sicher, dass die Buchführung der Unternehmung den geltenden Anforderungen entspricht. Es ist ein Inventar der zu diesem Zweck durchgeführten **Kontrollen**. Das IKS eignet sich jedoch auch hervorragend, die regelmässigen Aktivitäten zu dokumentieren, welche unternommen werden, um direkte oder indirekte Schäden durch falschen KI-Einsatz oder durch Cyberangriffe nach Möglichkeit zu vermeiden.

Das IKS als Inventar der Kontrollen nennt dabei die jeweilige Massnahme und deren Durchführungsmodalität – hier zwei Beispiele: Bevor ein neuer institutionalisierter KI-Einsatz im Betrieb in Produktion geht, sollen ausgewählte Funktionen (Personen) den Einsatz prüfen und dann ggf. freigeben. Oder einmal pro Jahr sollen alle erteilten **Zugriffsberechtigungen** kritisch hinterfragt und bestätigt oder entzogen werden. Je nach Betriebsgrösse, -komplexität und zwingenden Anforderungen ist die Granularität des IKS individuell festzulegen. In jedem Fall gilt jedoch, dass das IKS (also die Gesamtheit der Massnahmen) handhabbar und ökonomisch sein soll.[22]

Handlungsempfehlungen für KMU-Führungskräfte

Die rasante Entwicklung von KI-Technologien und die zunehmenden Bedrohungen durch Cyberangriffe erfordern von Führungskräften ein proaktives und umfassendes Handeln. Um ihrer Pflicht in Bezug auf KI und Cybersicherheit nachzukommen, sollten Führungskräfte folgende Massnahmen ergreifen:

1. **Einführung eines themenspezifischen Weisungswesens für KI und Cybersicherheit:** Führungskräfte müssen klare Richtlinien für die Implementierung und Nutzung von KI-Technologien festlegen. Dies umfasst die Festlegung von Verantwortlichkeiten, die Definition von Prozessen zur Überwachung und Bewertung der KI-Anwendungen sowie regelmässige Audits zur Sicherstellung der Einhaltung der Vorgaben. Schulungen der Mitarbeiter:innen in diesen Richtlinien sind unerlässlich, um eine korrekte und verantwortungsvolle Nutzung der KI-Systeme zu gewährleisten. Für die Cybersicherheit ist ein passendes Kontrollrahmenwerk (Framework, bspw. NIST CSF 2.0) einzuführen.
2. **Transparenz durch Offenlegung der KI-Verwendung:** Unternehmen sollten klare Richtlinien für die Offenlegung der KI-Verwendung erstellen. Nutzer:innen müssen informiert werden, wann und wie KI-Technologien eingesetzt werden und wie Daten gesammelt und verarbeitet werden. Eine transparente Kommunikation fördert das Vertrauen der Nutzer:innen und hilft, ethische Bedenken zu adressieren.
3. **Sicherstellung der Qualität des KI-Outputs:** Mechanismen zur kontinuierlichen Überprüfung und Validierung der KI-Modelle sind notwendig, um die Zuverlässigkeit und Genauigkeit der Ergebnisse zu gewährleisten. Regelmässige Updates und Leistungsüberprüfungen der Algorithmen sowie die Einbindung menschlicher Expert:innen in den Überprüfungsprozess tragen zur Vermeidung von Verzerrungen oder Fehlern bei.
4. **Etablierung eines internen Kontrollsystems (IKS):** Massnahmen im Zusammenhang mit der Sicherheit oder der Compliance in den Bereichen KI und Cybersicherheit sollten zentral im Rahmen des IKS dokumentiert werden. Das IKS erfasst die durchgeführten Kontrollen und dient der Dokumentation regelmässiger Aktivitäten zur Vermeidung von Schäden durch falschen KI-Einsatz oder Cyberangriffe. Die Granularität des IKS sollte an die Betriebsgrösse und -komplexität angepasst sein.

22 Bund (CH), Implementierung des Risikomanagementkonzepts (KMU, www.kmu.admin.ch/kmu/de/home/praktisches-wissen/finanzielles/risikomanagement/wie-fuehrt-man-ein-risikomanagementsystem-ein/implementierung.html).

Checkliste für Führungskräfte bezüglich KI und Cybersicherheit:

- ☐ Richtlinien für den Einsatz von KI-Technologien und Cybersicherheit festlegen
- ☐ Verantwortlichkeiten und Überwachungsprozesse für beide Themen definieren
- ☐ Regelmässige Audits und Schulungen für Mitarbeiter:innen durchführen
- ☐ Transparente Kommunikation und Offenlegung der KI-Verwendung sicherstellen
- ☐ Mechanismen zur Qualitätssicherung des KI-Outputs implementieren
- ☐ Menschliche Expert:innen in den Überprüfungsprozess von KI einbinden
- ☐ Technische und organisatorische Cybersicherheitsmassnahmen (TOM) implementieren
- ☐ Ein internes Kontrollsystem (IKS) zur Dokumentation und Überwachung etablieren
- ☐ Regelmässige Updates und Leistungsüberprüfungen der Algorithmen durchführen[23]
- ☐ Allfällige weitere, situationsbezogene Massnahmen (firmenspezifisch)

Durch die Umsetzung dieser Massnahmen und die Verwendung der Checkliste können Führungskräfte sicherstellen, dass die Nutzung von KI-Technologien und der Schutz vor Cyberbedrohungen verantwortungsvoll und effektiv gehandhabt werden. Dies trägt nicht nur zur Sicherheit und Compliance bei, sondern stärkt auch das Vertrauen der Nutzer:inne und die Effektivität der eingesetzten Systeme.

Fazit

KI ist eine mächtige Waffe – für Angreifer:innen und Verteidiger:innen.

Führungskräfte müssen sich bewusst sein, dass KI-basierte Cyberangriffe zunehmen, gleichzeitig werden aber auch KI-gestützte Schutzmassnahmen immer besser. Entscheidend ist es, die Risiken zu verstehen, die richtigen Sicherheitsstrategien zu etablieren und Mitarbeiter:innen entsprechend zu schulen. Nur so kann KI in der Cybersecurity sicher und verantwortungsvoll genutzt werden. Zudem ist auch die Sicherheit der KI-Lösung selbst zu gewährleisten. Hierzu gehört deren Qualitäts-Check und angemessene organisatorische Einbettung.

To-Do:

Checkliste für Führungskräfte bezüglich KI und Cybersicherheit durcharbeiten

23 Bei Einsatz von firmenspezifischen KI-Lösungen oder beim Einsatz von sogenannten KI-Agenten (Agentive A.I.).