GLOBAL NETWORK & INTELLIGENCE

CII PRAXIS UPDATE

Risiken, Strategien und Lösungsansätze:

Deepfakes im Kontext von Cybersecurity & Compliance

Dr. Christopher Nehring

CYBER|NTELLIGENCE
.Institute

Inhalt

Summary 3
Teil 1: Deepfakes im Schnellüberblick 5
Teil 2: Arten von Bedrohungen 7
Teil 3: Gegen- und Abwehrmaßnahmen 10
Teil 4: CII-Checkliste 14

Über den Autor

Dr. Christopher Nehring
Intelligence Director I cyberintelligence.institute

Dr. Nehring ist Forscher, Analyst & Experte für zahlreiche Medien in Deutschland und Europa. Sein Forschungsgebiet umfasst die Arbeit von Geheimdiensten, Desinformation, Hybride Kriegsführung & KI-Cyberrisken, insbesondere Deepfakes und Manipulationen. Er war Gastdozent & Experte für Desinformation des Medienprogramms der Konrad-Adenauer- Stiftung, wissenschaftlicher Leiter des Spionagemuseums Berlin sowie Senior Analyst des Institute for Global Analysis. Dr. Nehring ist regelmäßiger Gastautor im renommierten Medien. Zudem ist er als Speaker, Trainer & Berater für Unternehmen, Stiftungen und Medienorganisationen tätig.

Kontakt: **christopher.nehring@cyberintelligence.institute**

Co-Autor Teil 3:

Michael Ringtunatus
Softwareentwickler & Cybersecurity-Experte

Ringtunatus verfügt über langjährige Erfahrung in der Entwicklung von KI-gestützten Sicherheitslösungen. Als Mitgründer des Softwareunternehmens DeepDetectAI spezialisiert er sich auf die Erkennung von Deepfakes und Identitätsbetrug. Ringtunatus ist Mitglied der CISO Allianz Deutschland.

Summary

Deepfakes bedeuten eine große Bandbreite an Risiken und Bedrohungen für Unternehmen: von Betrug über Fälschungen bis hin zu gezielten Image-Angriffen. Unternehmen sind durch Deepfakes nicht nur technisch, sondern auch juristisch und reputationsbezogen verwundbar. Um die rasante technologische Entwicklung und die gesamte Bandbreite der Angriffsvektoren abzudecken, braucht es spezialisiertes Gefahrenbewusstsein und eine ganzheitliche Sicherheitsstrategie.

Heute schon an morgen denken

Die Bandbreite und Anzahl von Deepfake-Bedrohungen werden durch die schnelle technologische Entwicklung in der nahen Zukunft weiter zunehmen. Die nächste Stufe der Bedrohung ist bereits sichtbar: KI-Agenten, die Deepfake-Angriffe vollautomatisiert und autonom erstellen. Unternehmen müssen ihre Sicherheitsstrategien anpassen und gezielt Deepfake-Abwehrmechanismen entwickeln und integrieren: Um Risiken zu minimieren, braucht es eine breite Palette an technischen und menschlichen Schutzmaßnahmen. Sensibilisierung und Schulung von Führungskräften und Mitarbeitern sind essenziell, um Deepfake-Bedrohungen frühzeitig zu erkennen und abzuwehren. Langfristig sind Investitionen in Cybersecurity-Software und die Qualifizierung von Personal unvermeidlich und günstiger als die Schäden durch erfolgreiche Angriffe.

Essentials im Umgang mit Deepfakes

- ✓ KI-Bildung und Threat Awareness über Deepfake-Gefahren bei Mitarbeitern und Führungskräften sind die Basis für alle Schutzmaßnahmen.
- ✓ Erkennungssoftware ist ein Baustein, aber keine vollständige Lösung. Sie muss in erweiterte Cybersecurity-Lösungen integriert werden.
- ✓ Organisationen brauchen ganzheitliche Deepfake-Sicherheitsstrategien, die alle Gefahren adressieren, Abwehrkapazitäten aufbauen, Prozeduren und Incident Response definieren und dabei sowohl den Faktor Mensch als auch Technologien berücksichtigen.

Gegenmaßnahmen

Menschliche Gegenmaßnahmen	Technologische Gegenmaßnahmen
✓ Feste Standardabläufe für Finanztransaktionen (inkl. Ausnahmen)	✓ Echtzeit Deepfake-Erkennungstools
✓ Übermäßige Informationen auf Homepage vermeiden	✓ Monitoring & Threat Intelligence: Welche Deepfake-Angriffe sind gerade im Trend?
✓ Gesonderte Identifizierungsmechanismen für Meetings mit besonderer Bedeutung	✓ Spezielle technologische Fernauthentifizierungsmechanismen für Meetings mit besonderer Bedeutung
✓ Rück-/Sicherheitsfragen	✓ Sicherheitssoftware für Bilder, Videos und Inhalte auf der eigenen Homepage („Protective Shield“)
✓ Auf Instinkte achten	✓ Koppelung an IT-Security (z. B. Rückverfolgung und Dokumentation)
✓ Reminder bzw. Sensibilisierung für Passwörter-Teilen & Links (auch bei täuschend echten Anrufen)	✓ XDR & SOAR-Tools: Automatisierte Erkennung & Reaktion bei verdächtigem Verhalten oder Anomalien
✓ Betriebsklima & Umgang	✓ Cyberversicherungen (technisch-organisatorische Absicherung)
✓ Sensibilisierung von Finanzabteilungen, Buchhaltung & C-Level-Führungskräften	

Teil 1
Deepfakes im Schnellüberblick

Sowohl die Anzahl erfolgreicher Deepfake-Betrugsfälle als auch die dabei entstandenen Gesamtschäden sind in den vergangenen zwei Jahren seit der breiten Verfügbarkeit von KI-Technologien rapide angestiegen. Sicherheitsforscher erwarten, dass dieser Trend anhalten wird. Zum einen steigt die Qualität von KI-Tools und Deepfakes kontinuierlich, zum anderen werden fertige Deepfakes oder sogar komplette Drehbücher samt technischer Infrastruktur für Deepfake-Betrug inzwischen als Dienstleistung in Online-Foren und Darknet-Marktplätzen gehandelt.

Die Cybersecurity-Firma Tethris veröffentlichte hierzu im März 2025 einen eigenen Threat-Intelligence-Report über Unternehmen, die vollautomatisierte Deepfake-Betrugsangebote anbieten. Zahlreiche Fälle von auf Betrug spezialisierten Call-Centern in Indien, Südostasien, Georgien oder auch der Ukraine unterstreichen das Ausmaß und die globale Dimension dieser Entwicklung. Da KI- und Deepfake-Tools mithilfe automatischer Übersetzung Sprachbarrieren überwinden, agieren Angreifer häufig tausende Kilometer entfernt von ihren Opfern.

Während die Verbreitung von Deepfake- und KI-Betrugstechnologien zunimmt, geraten vor allem Einzelpersonen sowie kleine und mittelständische Unternehmen (KMU) zunehmend unter Druck, da sie mit den steigenden Cybersecurity-Anforderungen nicht mehr Schritt halten können. Infolgedessen sind nicht mehr nur Großunternehmen und globale Konzerne betroffen, sondern immer häufiger auch KMU.

Auf der Seite der Verteidigung entstehen parallel neue technologische Hilfsmittel und KI-gestützte Abwehrlösungen.

Deepfake-Erkennungstools werden in Kürze in gängige Video-Chat-Anwendungen wie Microsoft Teams oder Zoom integriert und perspektivisch auch in Messenger-Dienste.

Schon jetzt zeigt sich ein hochdynamischer und unübersichtlicher Anbietermarkt. Ähnliches gilt für KI-unterstützte XDR-, SOAR-, SIEM- und andere Vorfallserkennungstools. Diese Schutzmechanismen stehen allerdings vor der Herausforderung, Deepfake-Elemente in Videos, Anrufen und Meetings eindeutig von legitimen KI-Funktionen zu unterscheiden. Zudem kann die Integration von Deepfake-Erkennung in gängige Tools dazu führen, dass Angreifer auf ungeschützte Messenger-Plattformen ausweichen.

Um die Tragweite dieser Entwicklung in ihrer ganzen Breite zu verstehen, ist zunächst eine präzise Begriffsbestimmung erforderlich. Nur wenn klar ist, was unter einem „Deepfake“ verstanden wird und welche Erscheinungsformen diese Technologie annehmen kann, lassen sich die damit verbundenen Risiken systematisch erfassen.

Was ist ein Deepfake?

Der Begriff „Deepfake“ geht auf einen gleichnamigen Nutzer der Online-Plattform Reddit zurück, der bereits 2017 KI-Software einsetzte, um die Gesichter von Prominenten, vor allem in pornografischen Videos, zu tauschen. Dies war eine frühe Form des KI-basierten „Face Swap“.

Heute ist „Deepfake“ (in der EU auch „Deep Fakes“) ein offiziell anerkannter Begriff. Das EU-KI-Gesetz definiert Deepfakes als „durch KI erzeugte oder manipulierte Bild-, Ton- oder Videoinhalte, die realen Personen, Objekten, Orten oder Ereignissen

ähneln und einer Person fälschlich als echt erscheinen würden“. Kurz gesagt: Deepfakes sind synthetische Medien, die mithilfe von KI produziert oder verändert werden und oft kaum von echten Aufnahmen zu unterscheiden sind.

Zur Erstellung von Deepfakes existieren zahlreiche Tools und Anwendungen, die verschiedene Arten von Inhalten erzeugen können, darunter:

- KI-generierte oder manipulierte Bilder
- KI-generierte oder manipulierte Videos
- Voice Clones (KI-generierte Stimmkopien als Audiodatei)
- Live Voice Clones (z.B. in Telefon-gesprächen)
- Face-Swap-Bilder
- Face-Swap-Videos
- Live Face Swap (Gesichtertausch in Videocalls oder Livestreams)
- Live Face Swap mit Voice Clone (gleichzeitiger Gesichts- und Stimmentausch in Echtzeit)

Die reine Definition verdeutlicht zwar die technische Grundlage, sie erklärt jedoch nicht die wachsende Relevanz für Wirtschaft und Gesellschaft. Erst der Blick auf Häufigkeit und Schadenshöhe zeigt, warum Deepfakes als ernstzunehmende Bedrohung eingestuft werden müssen.

Risiken & Gefahren von Deepfakes

Die Häufigkeit und die verursachten Schäden durch Deepfakes nehmen in rasantem Tempo zu. Zwischen 2017 und 2022 wurden lediglich 22 Vorfälle registriert. Im Jahr 2023 verdoppelte sich die Zahl fast auf 42 und stieg 2024 auf 150. Bereits im ersten Quartal 2025 wurden mindestens 179 bekannte Fälle gemeldet, was eine exponentielle Zunahme verdeutlicht.

Parallel dazu wächst der finanzielle Schaden erheblich. Das Cybersecurity-Unternehmen Regularforensics ermittelte für deutsche Unternehmen im Jahr 2024 durchschnittliche Verluste von über 300.000 US-Dollar pro Vorfall. Deloitte bezifferte den globalen Schaden durch Deepfakes 2023 auf 12,3 Milliarden US-Dollar und prognostizierte bis 2027 einen Anstieg auf bis zu 40 Milliarden US-Dollar.

Auch Strafverfolgungsbehörden bewerten die Bedrohung zunehmend kritisch. So warnte Europol bereits 2022, dass Deepfakes insbesondere beim CEO-Fraud zum Standardwerkzeug werden könnten, eine Einschätzung, die sich seither bestätigt hat. Das FBI stuft die Technologie ebenfalls als akute Gefahr für Unternehmen aller Größenordnungen ein.

Eine von Deloitte im Jahr 2024 durchgeführte Befragung unter internationalen Führungskräften unterstreicht diese Einschätzung: Fast 26 Prozent der Befragten gaben an, im Vorjahr Opfer eines Deepfake-Betrugs geworden zu sein, während 52 Prozent von einer weiteren Zunahme solcher Angriffe ausgehen. Laut Sumsup stieg die Zahl der Deepfake-Angriffe 2024 in Deutschland um rund 145 Prozent, in den USA sogar um mehr als 300 Prozent. Besonders betroffen sind Finanzdienstleister, Banken, Versicherungen, Kommunikationsanbieter, Online-Marktplätze, Händler sowie die Gaming-Industrie.

Teil 2
Arten von Bedrohungen

Deepfakes heben Informationsangriffe, negative PR und Schmierkampagnen gegen Unternehmen sowie gegen Einzelpersonen auf eine neue Stufe. Deepfakes ermöglichen die glaubhafte Nachahmung von Führungskräften und erweitern klassische Methoden wie Phishing, Social Engineering und Cyberangriffe. Die Bandbreite reicht von gefälschten Unternehmensbilanzen über manipulierte Produktbewertungen bis hin zu rassistischen oder beleidigenden Äußerungen. Jede Form der Negativdarstellung oder Diffamierung kann mit täuschend echt wirkenden, emotional aufgeladenen und scheinbar objektiven „Beweisen“ untermalt werden.

Experten gehen davon aus, dass Deepfake-Technologie zunehmend auch zur Erstellung gefälschter Beweise für Erpressungen eingesetzt wird, etwa in Form kompromittierender Videos oder manipulierter Screenshots sensibler Daten. Die nachfolgenden Unterabschnitte differenzieren zentrale Bedrohungsarten, die durch Deepfake-Technologien in besonderem Maße begünstigt oder verstärkt werden.

Whaling

Whaling ist eine spezielle Form von Spear-Phishing-Angriffen, bei denen gezielt Führungskräfte ins Visier genommen werden. Diese Attacken beginnen häufig per E-Mail, erfolgen jedoch zunehmend direkt über Deepfakes am Telefon oder in Video-Calls. Typische Szenarien sind Anweisungen zu Zahlungen und Transaktionen (bekannt als CEO-Fraud) oder die Sabotage von Geschäftsprozessen durch falsche Arbeitsaufträge. Deepfake-Angreifer stören wichtige Abläufe, indem sie in Meetings als leitende Mitarbeiter auftreten oder über Telefon und Video-Calls täuschend echt wirkende Anweisungen erteilen. Darüber hinaus nutzen Cyberkriminelle diese Taktik, um Passwörter abzugreifen oder Mitarbeitende dazu zu bewegen, auf infizierte Links zu klicken.

CEO-Fraud

CEO-Fraud bezeichnet eine Betrugsmasche, bei der sich Angreifer als hochrangige Führungskraft ausgeben, um Mitarbeitende zu Überweisungen hoher Geldbeträge oder zu anderen schädlichen Handlungen zu bewegen. Diese Methode ist nicht neu, wird jedoch durch Deepfake-Technologie erheblich erleichtert. Oft beginnt der Angriff mit einer Spear-Phishing-E-Mail, in der ein angeblicher CEO dringende und geheime Geschäftsvorgänge vortäuscht. Beliebte Vorwände sind geplante Übernahmen, Steuermodelle oder Abfindungen. In vielen Fällen werden die Opfer anschließend durch Anrufe mit Deepfake-Stimmen oder in Video-Calls zusätzlich unter Druck gesetzt. Dringlichkeit und Geheimhaltung sind Kernelemente dieser Angriffe. Täter setzen die Opfer psychologisch unter Druck, indem sie Autorität, Drohungen, aber auch Empathie oder Belohnungen simulieren. Deepfakes verstärken die Glaubwürdigkeit und schaffen eine scheinbar stimmige Gesamtsituation. Die Qualität der Angriffe reicht von plumpen Versuchen bis hin zu professionell orchestrierten Täuschungen. Ein einfaches Beispiel war der gescheiterte Deepfake-Fraud bei Ferrari im Jahr 2024. Ein Manager erhielt WhatsApp-Nachrichten angeblich vom CEO, durchschaute den Betrug jedoch durch eine gezielte Rückfrage. Ein hochprofessionelles Beispiel ereignete sich bei der britischen Baufirma Arup in Hongkong im Januar 2024. Dort wurde ein leitender Mitarbeiter durch ein Video-Meeting mit mehreren täuschend echt gefälschten Chefs zur Autorisierung einer Überweisung von 25 Millionen US-Dollar bewegt.

Pig Butchering Scam

Diese Betrugsmasche basiert auf falschen Versprechungen lukrativer Investitionen, häufig im Bereich Kryptowährungen. Täter bauen über Wochen oder Monate Vertrauen auf, zeigen anfänglich kleine Gewinne und bewegen die Opfer schließlich zu hohen Investitionen. Danach verschwinden sie mit den Geldern. Deepfakes kommen hier immer häufiger zum Einsatz.
In Georgien wurde 2025 ein Callcenter mit über 80 Mitarbeitenden aufgedeckt, das seit 2022 über 6000 Opfer um rund 35 Millionen US-Dollar betrogen haben soll. Dabei nutzten die Täter Deepfake-Videos von Prominenten zur Bewerbung angeblicher Produkte oder setzten Live-Deepfakes bei Telefon- und Video-Calls ein. In einem anderen Fall wurde ein Galerist in Nordengland über Monate von einem Deepfake des Schauspielers Pierce Brosnan getäuscht.

Romance Scam

Beim Romance Scam bauen Betrüger emotionale Bindungen zu ihren Opfern auf, meist über Dating-Plattformen oder soziale Netzwerke. Sobald eine Abhängigkeit entsteht, täuschen sie Notsituationen wie medizinische Notfälle oder Reiseprobleme vor und erbitten Geldüberweisungen. Neben finanziellen Verlusten entstehen für die Opfer oftmals erhebliche emotionale Schäden. Dank Deepfakes hat diese Betrugsmasche inzwischen industrielle Dimensionen erreicht.
Die nigerianische Bande der Yahoo-Boys setzt Live-Deepfakes ein und verteilt Anleitungen in Chat-Räumen. Plattformen wie Haotian.AI in Kambodscha haben den Prozess sogar weitgehend automatisiert, von der Kontaktaufnahme durch Chatbots bis hin zu Deepfake-Personas für Live-Gespräche.

Cybermobbing

Deepfakes werden gezielt zur Beleidigung, Rufschädigung und Herabwürdigung eingesetzt. Besonders verbreitet ist sogenannter „Deep Porn“, bei dem Gesichter realer Personen in pornografische Inhalte montiert werden. Prominente, Politikerinnen und zunehmend auch Männer sind davon betroffen. Ein bekanntes Beispiel war die Sängerin Taylor Swift, deren Name auf der Plattform X zeitweise gar nicht mehr gesucht werden konnte, weil so viele Deepfake-Pornos kursierten. Auch andere Formen von Mobbing nehmen zu. In den USA wurden Schulleiter mit Deepfake-Videos diffamiert, die ihnen rassistische Äußerungen unterstellten. In mehreren Fällen wurden Schüler oder Lehrer als Täter identifiziert.

Produktpiraterie und -fälschungen

Deepfakes werden für die visuelle Täuschung bei Produktfälschungen eingesetzt, insbesondere auf Online-Marktplätzen. Ein Beispiel betrifft die US-Firma Mattel, deren Puppenmarke durch KI-generierte Bilder einer schwangeren Ken-Puppe in Verruf gebracht wurde. Das Problem verschärft sich dadurch, dass auch seriöse Unternehmen zunehmend KI-Inhalte einsetzen, wodurch die Grenzen zwischen authentischen und gefälschten Darstellungen verschwimmen. Viele Plattformen verfügen bislang über keine wirksamen Mechanismen zum Schutz vor KI-generierten Inhalten.

Identitätsfälschung

Deepfake-Technologie eignet sich in besonderem Maße zur Fälschung von Identitäten. Russische Händler bieten beispielsweise KI-generierte Ausweise westlicher Staaten an, die für Online-Käufe, Fake-Accounts oder den Handel mit Kryptowährungen genutzt werden. Auch Bewerbungsprozesse sind betroffen. In einer Umfrage gaben 17 Prozent der HR-Manager an, bereits Deepfake-Bewerber erlebt zu haben. Nordkorea nutzt diese Methode gezielt, um IT-Fachkräfte unter falscher Identität in westlichen Unternehmen unterzubringen. So können Gehälter abgeschöpft und zugleich Spione platziert werden. Im Februar 2025 flogen zwei Deepfake-Bewerber bei der Firma Vidoc Security Lab in Polen auf, nachdem sie fast den gesamten Bewerbungs-prozess durchlaufen hatten.

Versicherungsbetrug

Versicherungen verzeichnen einen deutlichen Anstieg an KI-basierten Betrugsversuchen. Täter reichen gefälschte Nachweise wie manipulierte Fotos, Rechnungen oder Gutachten ein. Besonders häufig sind fingierte Kfz-Schäden, aber auch Immobilien- und Diebstahlversicherungen sind betroffen. In einem Fall wurde sogar ein KI-generiertes Überwachungsvideo als angeblicher Nachweis für eine Verletzung eingereicht.

Nachweis- und Beweisfälschung

Deepfakes machen es einfach, kostengünstig und schnell Beweise wie Quittungen, Rechnungen, Protokolle oder Berichte zu fälschen. Auf diese Weise eröffnen sich für Betrüger zahlreiche Möglichkeiten, Unternehmen, Behörden und Privatpersonen zu täuschen und finanziell zu schädigen.

Aktienkursmanipulation

Die Manipulation von Aktienkursen durch Falschmeldungen ist nicht neu, wird durch KI und Deepfakes aber erheblich erleichtert. Beispiele sind die gefälschte Übernahme-Meldung bei Twitter im Jahr 2015, die den Kurs um acht Prozent steigen ließ, sowie der Hack des offiziellen SEC-Kontos im Januar 2024, der massive Kursbewegungen bei Bitcoin auslöste.
Ein besonders prägnantes Beispiel ereignete sich 2023, als ein KI-generiertes Bild eines angeblichen Anschlags auf das Pentagon verbreitet wurde. Der S&P 500 verlor daraufhin kurzfristig fast 0,3 Prozent. Auch nach der US-Wahl 2024 kursierten zahlreiche Deepfakes über Elon Musk und sein Firmenimperium, die jeweils zu Kursbewegungen führten.

Beweislastumkehr und Deepfake-Defence

Die bloße Existenz von Deepfake-Technologie reicht aus, um Beweise in Zweifel zu ziehen. Vor Gericht wird dies gezielt genutzt. Ein Beispiel ist ein Verfahren zu Teslas Autopilot-Programm. Dort stellte die Verteidigung die Möglichkeit in den Raum, dass eine kritische Aussage von Elon Musk ein Deepfake sein könnte. Konkrete Beweise wurden nicht vorgelegt, doch allein der Hinweis auf die theoretische Möglichkeit reichte aus, um die Echtheit der Aussage infrage zu stellen und die Beweislast zu verschieben.

Teil 3
Gegen- & Abwehrmaßnahmen

Deepfakes stellen Unternehmen vor komplexe Herausforderungen, die technologische, organisatorische und rechtliche Dimensionen betreffen. Es gibt keine einzelne Gegenmaßnahme gegen Deepfake-Angriffe. Vielmehr existieren viele unterschiedliche Angriffsformen, die jeweils verschiedene Strategien & Sicherheitsmaßnahmen erfordern. Im Folgenden wird dargestellt, wie sich Organisationen wirksam schützen können.

Gängige Gegenmaßnahmen

- ✓ Awareness, Training, Resilienz („Human Firewall“)
- ✓ Etablierte Mechanismen für Incident Response und Reaktion
- ✓ Persönliche Identifizierung & Sicherheitspasswörter oder -phrasen
- ✓ Erweiterte Multi-Faktor-Authentifizierung, Identity- und Access-Management
- ✓ Integrierte Deepfake-Erkennungssoftware
- ✓ Frühwarnsysteme und Threat Intelligence
- ✓ Erweiterte Cybersecurity (z. B. SIEM, XDR, SOAR, Access Management, KI-basierte Spam- & Phishing-Erkennung)
- ✓ Digital Listening
- ✓ OSINT
- ✓ Informationsaustausch mit Experten, Forschern, Behörden, Medien sowie Public-Private-Partnerships & Kooperationen
- ✓ Interner abteilungsübergreifender Informationsaustausch
- ✓ Melden & Flaggen von Deepfake-Inhalten (koordiniert)
- ✓ Rechtliche Schritte: Reporting, Take-Downs, Anzeige & Strafverfolgung
- ✓ Strategische Kommunikation
- ✓ Crisis Response, Krisenmanagement & Krisenkommunikation
- ✓ Umfassende Dokumentation (z.B. Screenshots, Aufnahmen, Metadaten, Profildaten)

Herausforderungen bei der Bekämpfung

Derzeit lassen sich drei zentrale Herausforderungen bei der Bekämpfung von Deepfake-Betrug beobachten. Erstens ist die Dunkelziffer weiterhin sehr hoch, und viele Vorfälle werden nicht an Strafverfolgungsbehörden gemeldet. Zweitens schreitet die technologische Entwicklung im Bereich künstlicher Intelligenz so schnell voran, dass die Angreifer im Vorteil sind. Weder Erkennungssoftware noch deren Integration in bestehende Cybersecurity-Tools sind bislang ausgereift. Drittens bestehen Defizite beim allgemeinen Stand der Cybersecurity, insbesondere bei Ausstattung und Awareness. Aussagen wie „das kann uns nicht passieren“ oder „wir sind bereits gut aufgestellt“ sind weiterhin weit verbreitet und gerade bei Opfern von Deepfake-Betrug zu finden.

Der Ansatz: Mensch und Technologie gemeinsam

Deepfake-Angriffe zielen vor allem auf menschliche Nutzer und ihre Psyche ab, um Betrug oder Manipulation zu erreichen. Deshalb handelt es sich nicht nur um ein technisches Risiko. Studien zeigen

jedoch eindeutig, dass KI-basierte Systeme in der Lage sind, Deepfakes zuverlässiger zu erkennen als Menschen. Angesichts der bevorstehenden Entwicklung sogenannter agentischer KI, die Deepfake-Angriffe weiter verstärken wird, sind Menschen auf technologische Unterstützung angewiesen. Die Frage lautet daher nicht, ob Mensch oder Technologie die bessere Lösung ist, sondern wie Technologie Menschen bestmöglich unterstützen kann. Menschliche Entscheidungsträger, Führungskräfte und Mitarbeiter benötigen sowohl digitale Kompetenzen als auch KI-spezifische Fähigkeiten. Gleichzeitig muss Software so gestaltet sein, dass sie diese Anwender wirksam unterstützt. Im Mittelpunkt steht somit ein human-zentrierter Ansatz, bei dem Menschen mit der passenden Technologie in die Lage versetzt werden, Deepfake-Angriffe zu erkennen und abzuwehren. Die sogenannte „human firewall“, also ein geschulter und KI-kompetenter Nutzer, ist dabei der wichtigste Baustein. Nur wenn Mitarbeiter Bedrohungen erkennen, Resilienz aufbauen und die ihnen zur Verfügung stehenden technischen Tools zielgerichtet einsetzen können, entsteht wirksamer Schutz.

Was Menschen gegen Deepfakes tun können

Menschen sind häufig das eigentliche Ziel von Deepfake-Angriffen, vor allem bei Betrugsfällen. Daher können sie sich nicht ausschließlich auf technische Tools verlassen, sondern müssen eigene Fähigkeiten und ein Bewusstsein für Gefahren entwickeln. Ein geschulter Mitarbeiter ist vorbereitet, erkennt potenzielle Angriffe und achtet auf typische Marker, die Deepfakes verraten können. Dazu gehören optische und akustische Auffälligkeiten sowie die Prüfung von Kontext und Metadaten. Organisationen sollten ihre Mitarbeiter und Führungskräfte in Verfahren schulen, die bei einem vermuteten Angriff greifen. Dazu gehören persönliche Erkennungsinformationen, festgelegte Meldeketten und definierte Ansprechpartner.

Wichtig ist zudem, dass technische Tools wie Erkennungssoftware und Cybersecurity-Systeme nutzerfreundlich und verständlich gestaltet sind. Nur dann können sie im Ernstfall wirksam eingesetzt werden. KI-Bildung, Schulungen zum Umgang mit Deepfakes sowie klare Prozesse für Incident Response gehören ebenso zur Sicherheitsstrategie wie erweiterte KI-basierte Tools.

Technologische Maßnahmen

Neben menschlicher Wachsamkeit und Schulung sind technologische Instrumente unverzichtbar, um Deepfake-Betrug wirksam zu erkennen, Bedrohungen zu analysieren und geeignete Gegenmaßnahmen einzuleiten.

Deepfake-Erkennungssoftware

Die Erkennung von Deepfakes ist eine zentrale Herausforderung. Forschungsergebnisse belegen, dass KI-basierte Systeme besser als Menschen darin sind, synthetische Inhalte zu identifizieren. So erreicht die von Intel und der Binghamton University entwickelte FakeCatcher-Technologie Erkennungsraten zwischen 91 und 96 Prozent. Eine Studie von iProov aus dem Jahr 2025 zeigte, dass nur 0,1 Prozent der menschlichen Teilnehmer alle Deepfakes korrekt identifizieren konnten, selbst wenn sie nach Fälschungen suchten. Erkennungssoftware analysiert Bilder, Videos oder Audiodateien auf typische Manipulationsmuster. Dazu gehören unnatürliches Blinzeln, asynchrone Lippenbewegungen oder Artefakte bei Schatten und Licht. Neuere Ansätze erkennen zudem physiologische Signale wie minimale Farbveränderungen der Haut, die Rückschlüsse auf den Herzschlag zulassen. Parallel entwickelt sich die Technologie der Deepfake-Generatoren weiter, sodass auch solche Merkmale zunehmend gefälscht werden können. Fortschritte gibt es hingegen bei der feingranularen Detektion, die selbst kleinste Veränderungen in Ton oder Bild aufspüren kann.

Der Vorteil dieser Software liegt in ihrer Skalierbarkeit. Sie kann große Mengen an Daten in kurzer Zeit prüfen und liefert so Hinweise für menschliche Entscheidungen. Gleichzeitig ist sie auf Trainingsdaten, Parameter und regelmäßige Updates angewiesen, da sich die Erkennungsmuster ständig ändern.

Für Unternehmen bedeutet das: Deepfake-Erkennungssoftware kann ein hilfreiches Warnsystem sein, sie ist jedoch allein keine Lösung. Sie muss in eine umfassende Sicherheitsarchitektur integriert werden, die auch Prozesse für die Reaktion auf erkannte Angriffe vorsieht.

Advanced and Extended Cybersecurity Software

Die Erkennung eines Deepfakes löst das Problem noch nicht. Entscheidend ist, wie Organisationen im nächsten Schritt reagieren. Dazu benötigen sie Systeme, die erkannte Bedrohungen automatisch weiterleiten und Reaktionsprozesse auslösen. Moderne Sicherheitsarchitekturen integrieren Deepfake-Erkennung in bestehende Cybersecurity-Systeme. Zentral sind Security Information and Event Management Systeme (SIEM), die sicherheitsrelevante Daten aus verschiedenen Quellen sammeln und analysieren. Werden diese Systeme mit Deepfake-Erkennung gekoppelt, können sie Bedrohungen wie bei klassischen Cyberangriffen sichtbar machen und Reaktionen einleiten. Noch weiter gehen Extended Detection and Response (XDR) und Security Orchestration, Automation and Response (SOAR). Diese Plattformen ermöglichen nicht nur die Erkennung, sondern auch die automatisierte Reaktion auf Deepfake-Angriffe. Sie erkennen Bedrohungen über verschiedene Systeme hinweg, etwa E-Mail, Endgeräte oder Netzwerke, und koordinieren Gegenmaßnahmen. Dazu gehören die Sperrung von Nutzerkonten oder die Eskalation an spezielle Teams. Auch Identity and Access Management Systeme (IAM) spielen einewichtige Rolle, da viele Angriffe über unzureichend geschützte Zugänge erfolgen. Wird IAM durch Privileged Access Management erweitert, erhalten besonders sensible Konten zusätzliche Sicherheitsstufen wie Mehrfaktor-Authentifizierung. Ergänzend tragen auch User and Entity Behavior Analytics (UEBA) sowie Endpoint Detection and Response (EDR) zur Abwehr bei, da sie auffälliges Verhalten oder verdächtige Aktivitäten erkennen.

Effektiver Schutz entsteht somit erst durch die Kombination von Erkennungssoftware mit Cybersecurity-Systemen wie SIEM, IAM, PAM, XDR oder SOAR. Diese Systeme müssen technisch kompatibel sein und in die organisatorischen Abläufe eingebettet werden. Unternehmen wiederum sollten darauf achten, dass ihre Software sowohl Erkennung als auch Reaktion abdeckt und gleichzeitig nutzerfreundlich bleibt.

Ausblick

KI-gestützte Systeme, sogenannte Agenten, die eigenständig handeln, lernen und Aufgaben ausführen können, stellen die nächste Entwicklungsstufe im Bereich der künstlichen Intelligenz dar. Dies betrifft auch Cybersecurity-Risiken und insbesondere Deepfakes. Angriffe lassen sich mit solchen Systemen nicht nur inhaltlich planen und vorbereiten, sondern auch direkt umsetzen. Zukünftig werden KI-Agenten Informationen aus offenen Quellen oder sozialen Netzwerken sammeln, auf dieser Basis Deepfakes erstellen und Angriffe eigenständig durchführen können. Dazu zählen automatisierte Anrufe, E-Mails oder die Teilnahme an Videokonferenzen. Die Gefahr solcher Angriffe liegt vor allem in ihrer hohen Skalierbarkeit: Ein einzelner Agent ist in der Lage, zahlreiche Deepfake-Angriffe gleichzeitig durchzuführen, sich dabei kontinuierlich zu verbessern und Abwehrstrategien anzupassen. Diese Entwicklung betrifft nicht nur die Angreifer, sondern auch die Cybersecurity, die ihrerseits auf KI-Agenten setzen muss.

Agentenbasierte Verteidigungssysteme können beispielsweise Kommunikation, Verhalten und Kontext in Echtzeit analysieren, um Angriffsstrategien zu erkennen, abzuwehren und automatisierte Gegenmaßnahmen einzuleiten. Zentrale Bausteine solcher Systeme sind adaptive Zugangskontrollen, Zero-Trust-Architekturen und eine erweiterte Rechteverwaltung. Adaptive Zugangskontrollen passen ihre Prüfmechanismen dynamisch an Kontextfaktoren wie Uhrzeit oder Standort an und schaffen dadurch gezielten Schutz. Zero-Trust-Architekturen gehen davon aus, dass kein Kontakt und kein Inhalt grundsätzlich vertrauenswürdig ist, weshalb jede Kommunikation überprüft und authentifiziert werden muss. Ergänzend sorgt eine konsequente Rechteverwaltung dafür, dass privilegierte Zugänge streng kontrolliert und gegen Manipulation gesichert sind.

Empfehlungen

Bei der Bekämpfung von Deepfake- und CEO-Betrug stehen sich häufig zwei Sichtweisen gegenüber. Hersteller von Erkennungssoftware betonen, dass nur KI-gestützte Systeme Deepfakes zuverlässig identifizieren können. Trainer und Schulungsanbieter verweisen hingegen auf den Faktor Mensch als zentrale Schwachstelle und plädieren für Qualifizierung und Upskilling von Mitarbeitern und Führungskräften.

Der Ansatz des CII setzt darauf, diese beiden Perspektiven miteinander zu verbinden. Ein wirksames Schutzkonzept kann nur durch eine Kombination von menschlichen und technologischen Fähigkeiten entstehen. Vorbereitung und Resilienzbildung sind dabei auf beiden Seiten von entscheidender Bedeutung. Mitarbeiter und Führungskräfte müssen regelmäßig für die Funktionsweise von Deepfake-Betrug sensibilisiert werden. Hierbei ist es notwendig, aktuelle Threat Intelligence zu berücksichtigen. Besonders geeignet sind Angriffs-Simulationen im Rahmen von Trainings und Übungen. Ebenso wichtig ist die allgemeine Cybersecurity-Awareness, etwa beim Umgang mit Links, E-Mails oder Passwörtern. Darüber hinaus spielt auch die Unternehmenskultur eine wichtige Rolle. In stark hierarchischen Organisationen befolgen Mitarbeiter Anweisungen vermeintlicher Führungskräfte oft ungeprüft, selbst wenn Zweifel bestehen. Deshalb sind feste Abläufe, Identifizierungsmechanismen, Codewörter oder Sicherheitsfragen wichtige Maßnahmen, die durch technische Lösungen ergänzt werden können. Übermäßig detaillierte Informationen über Führungspersonal auf Unternehmenswebseiten oder ungesicherte Foto- und Videodateien erhöhen zudem die Angriffsfläche für Betrüger. Auf technischer Ebene ist die Kombination verschiedener Tools entscheidend. Klassische Cybersecurity-Systeme wie SIEM, XDR und SOAR können Deepfakes in Videokonferenzen nicht eigenständig erkennen. Spezialisierte Erkennungssoftware hingegen identifiziert zwar die Manipulation, liefert jedoch keine Handlungsanweisungen und kann Angriffe nicht stoppen. Erst die Kombination dieser Systeme ermöglicht einen wirksamen Schutz. Hinzu kommt, dass Softwarelösungen auf die Bedürfnisse der Nutzer abgestimmt sein müssen. Nur wenn Mitarbeiter die Systeme verstehen, navigieren und im Ernstfall effektiv bedienen können, entfalten sie ihre Schutzwirkung.

Teil 4
CII-Checkliste

Angriffsart	Menschliche Maßnahmen	Technologische Maßnahmen
Betrugsangriffe CEO-Fraud, Pig Butchering & Romance Scam	✓ Awareness, Training, Resilienz ✓ Sicherheitsphrasen ✓ interne Abläufe & Freigaben ✓ interner & externer Informationsaustausch	✓ MFA, IAM ✓ Deepfake-Erkennung ✓ erweiterte Cybersecurity ✓ Threat Intelligence
Whaling Angriffe auf Führungskräfte	✓ Sensibilisierung von Executives ✓ klare Entscheidungswege ✓ Crisis Response	✓ erweiterte MFA & IAM ✓ Deepfake-Erkennung ✓ Threat Intelligence ✓ SIEM, XDR, SOAR ✓ Anbindung an Behörden
Reputations- & Imageangriffe	✓ Awareness & Medientraining ✓ strategische Krisenkommunikation	✓ Deepfake-Erkennung ✓ Digital Listening ✓ OSINT ✓ Austausch mit Medien & Behörden ✓ Flagging & Takedowns
Cybermobbing	✓ Awareness bei Jugendlichen, Eltern & Mitarbeitenden ✓ Kommunikation ✓ Dokumentation	✓ Deepfake-Erkennung ✓ Digital Listening ✓ OSINT ✓ Melden & Flagging ✓ Plattform-Takedowns
Produktpiraterie & Fälschungen	✓ Kooperation mit Partnern & Behörden ✓ Sensibilisierung	✓ Deepfake-Erkennung ✓ Cybersecurity mit Tracking ✓ Digital Listening ✓ OSINT ✓ Takedowns
Identitätsfälschung	✓ Awareness & Aufklärung ✓ Sicherheitsphrasen ✓ Meldung & Anzeige	✓ MFA & IAM ✓ Deepfake-Erkennung ✓ Threat Intelligence ✓ Dokumentation
Nachweis- & Beweisfälschung	✓ juristische Bewertung ✓ strategische Kommunikation	✓ Deepfake-Erkennung ✓ Threat Intelligence ✓ erweiterte Forensik & Logging
Versicherungsbetrug	✓ Austausch mit Versicherern & Ermittlern	✓ Deepfake-Erkennung ✓ Threat Intelligence ✓ erweiterte Cybersecurity
Aktienkursmanipulation	✓ strategische Krisenkommunikation	✓ Deepfake-Erkennung ✓ Threat Intelligence ✓ Digital Listening ✓ Austausch mit Finanzaufsicht
Beweislastumkehr „Deepfake Defence“	✓ Awareness zu Deepfake-Taktiken ✓ Crisis Response ✓ lückenlose Dokumentation	✓ Deepfake-Erkennung ✓ juristische & technische Forensik

Literaturverzeichnis

Teil 1

Europäische Union (2025): Artificial Intelligence Act, Artikel 3 (60). In: Artificial Intelligence Act – Definitionen. Verfügbar unter: https://artificialintelligenceact.eu/article/3/ [Abruf: 13.01.2025].

Teil 2

Coinpaper (2025): AI-generated fake IDs challenge crypto exchanges. Verfügbar unter: https://coinpaper.com/3251/ai-generated-fake-i-ds-challenge-crypto-exchanges-kyc-protocols [Abruf: 21.04.2025].

CNN (2025): Arup deepfake scam loss in Hong Kong. Verfügbar unter: https://edition.cnn.com/2024/05/16/tech/arup-deepfake-scam-loss-hong-kong-intl-hnk/index.html [Abruf: 21.04.2025].

Crowdstrike (2025): Global Threat Report 2025. Verfügbar unter: https://go.crowdstrike.com/2025-global-threat-report.html [Abruf: 18.06.2025].

Daily Mail (2025): AI deepfake 007 Pierce Brosnan art gallery. Verfügbar unter: https://www.dailymail.co.uk/news/article-14376127/AI-deepfake-007-Pierce-Brosnan-art-gallery.html [Abruf: 21.04.2025].

Deloitte (2025a): Deepfake banking fraud risk on the rise. Verfügbar unter: https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk- on-the-rise.html [Abruf: 18.06.2025].

Deloitte (2025b): Generative AI and the fight for trust. Verfügbar unte: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/Advisory/us-generative-ai-and-the-fight-for-trust.pdf[Abruf: 18.06.2025].

Europol (2025): Deepfake technology could become staple tool for organised crime. Verfügbar unter: https://www.europol.europa.eu/media-press/newsroom/news/europol-report-finds-deepfake-technology-could-become- staple-tool-for-organised-crime [Abruf: 21.04.2025].

FBI (2025): Business Email Compromise. Verfügbar unter: https://www.fbi.gov/how-we-can-help-you/scams-and-safety/common-frauds-and-scams/business-email-compromise [Abruf: 18.06.2025].

Golem (2025): Ferrari entgeht raffiniertem KI-basierten Identitätsbetrug. Verfügbar unter: https://www.golem.de/news/ceo-nachgeahmt-ferrari-entgeht-raffiniertem-ki-basierten- identitaetsbetrug-2407-187506.html[Abruf: 21.04.2025].

Independent (2025): Taylor Swift deepfake case. Verfügbar unter: https://www.independent.co.uk/news/new-york-ap-taylor-swift-swifties-facebook-b2485673.html [Abruf: 18.06.2025].

Kennedys Law (2025): Deepfakes in the insurance market. Verfügbar unter: https://kennedyslaw.com/en/thought-leadership/article/2024/deepfakes-in-the-insurance-market-a-personal-injury-perspective/ [Abruf: 18.06.2025].

Lepide (2025): Ransomware and deepfake technology. Verfügbar unter: https://www.lepide.com/blog/ransomware-and-deepfake-technology/ [Abruf: 18.06.2025].

Misbar (2025): Pregnant Ken doll claim. Verfügbar unter: https://misbar.com/en/factcheck/2022/06/03/pregnant-ken-doll-claim-originated-from-satirical-website [Abruf: 21.04.2025].

Reality Defender (2025): Deepfake Whaling Defense. Verfügbar unter: https://www.realitydefender.com/blog/deepfake-whaling-defense [Abruf: 21.04.2025].

Regula Forensics (2025): Deepfake fraud costs. Verfügbar unter: https://regulaforensics.com/news/deepfake-fraud-costs/[Abruf: 18.06.2025].

Resume Genius (2025): AI impact on hiring. Verfügbar unter: https://resumegenius.com/blog/job-hunting/ai-impact-on-hiring [Abruf: 21.04.2025].

Spiegel (2025): Taylor Swift: X unterbindet Suche nach Fake-Pornomotiven. Verfügbar unter: https://www.spiegel.de/netzwelt/netzpolitik/taylor-swift-x-unterbindet-suchen-nach-fake-pornomotiven-a-d8ddb5fe-94af-40a0-982a-d008f3ed4fe8 [Abruf: 21.04.2025].

Sumsub (2025): Deepfake cases surge in countries holding 2024 elections. Verfügbar unter: https://sumsub.com/newsroom/deepfake-cases-surge-in-countries-holding-2024-elections-sumsub-research-shows/ [Abruf: 21.04.2025].

Tehtris (2025): Deepfake-as-a-Service Threat Intelligence Report. Verfügbar unter: https://www.datensicherheit.de/wp-content/uploads/tehtris-deepfake-as-a-service-threat-intelligence-report.pdf [Abruf: 18.06.2025].

The Fashion Law (2025): Williams-Sonoma takes on Dupe.com in lawsuit over dupe culture. Verfügbar unter: https://www.thefashionlaw.com/williams-sonoma-takes-on-dupe-com-in-lawsuit-over-dupe-culture/ [Abruf: 18.06.2025].

The Guardian (2025a): Deepfakes, cash and crypto: how call centre scammers duped 6000 people. Verfügbar unter: https://www.theguardian.com/money/2025/mar/05/deepfakes-cash-and-crypto-how-call-centre- scammers-duped-6000-people [Abruf: 18.06.2025].

The Guardian (2025b): Elon Musks statements could be deepfakes, Tesla defence lawyers tell court. Verfügbar unter: https://www.theguardian.com/technology/2023/apr/27/elon-musks-statements-could-be- deepfakes-tesla-defence-lawyers-tell-court [Abruf: 18.06.2025].

The Register (2025): IT worker scam. Verfügbar unter: https://www.theregister.com/2025/02/11/it_worker_scam/ [Abruf: 18.06.2025].

US Treasury (2025): Managing Artificial Intelligence-Specific Cybersecurity Risks in the Financial Services Sector. Verfügbar unter: https://home.treasury.gov/system/files/136/Managing-Artificial-Intelligence-Specific- Cybersecurity-Risks-In-The-Financial-Services-Sector.pdf [Abruf: 21.04.2025].

Wired (2025): Yahoo boys and real-time deepfake scams. Verfügbar unter: https://www.wired.com/story/yahoo-boys-real-time-deepfake-scams/ [Abruf: 21.04.2025].

Zurich Insurance (2025): Insurance must prepare for rise in deepfake AI fraud. Verfügbar unter: https://www.zurich.co.uk/news-and-insight/insurance-must-prepare-for-a-rise-in-deepfake-ai-fraud [Abruf: 21.04.2025].

Teil 3

Computer Weekly (2025): SOAR – Security Orchestration, Automation and Response. Verfügbar unter: https://www.computerweekly.com/de/definition/SOAR-Security-Orchestration-Automation-and-Response [Abruf: 16.07.2025].

Cyberark (2025): Privileged Access Management. Verfügbar unter: https://www.cyberark.com/de/what-is/privileged-access-management/ [Abruf: 16.07.2025].

Deloitte (2025): Deepfake banking fraud risk on the rise. Verfügbar unter: https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud- risk-on-the-rise.html [Abruf: 02.08.2025].

Enginsight (2025): SIEM. Verfügbar unter: https://enginsight.com/de/glossar/siem/ [Abruf: 02.08.2025].

IDW (2025): Deepfakes now have a heartbeat. Verfügbar unter: https://nachrichten.idw-online.de/2025/05/20/deepfakes-now-have-a-heartbeat-challenging-current-detection-methods [Abruf: 16.07.2025].

iProov (2025): Study reveals deepfake blindspot. Verfügbar unter: https://www.iproov.com/de/press/study-reveals-deepfake-blindspot-detect-ai-generated-content [Abruf: 16.07.2025].

McDuff, D. et al. (2025): Remote Photoplethysmography (rPPG). In: PMC. Verfügbar unter: https://pmc.ncbi.nlm.nih.gov/articles/PMC9267568/ [Abruf: 02.08.2025].

Microsoft (2025a): Identity & Access Management (IAM). Verfügbar unter: https://www.microsoft.com/de-de/security/business/security-101/what-is-identity-access-management-iam [Abruf: 16.07.2025].

Microsoft (2025b): Endpoint Detection & Response (EDR). Verfügbar unter: https://www.microsoft.com/de-de/security/business/security-101/what-is-edr-endpoint-detection-response [Abruf: 02.08.2025].

Nehring, C. (2025): How to Recognize Deepfakes?. Verfügbar unter: https://www.linkedin.com/in/christopher-n-423b06257/overlay/1714981524276/single-media-viewer [Abruf: 02.08.2025].

Palo Alto Networks (2025): Extended Detection and Response (XDR). Verfügbar unter: https://www.paloaltonetworks.de/cyberpedia/what-is-extended-detection-response-XDR [Abruf: 16.07.2025].

Regula Forensics (2025): Deepfake fraud costs. Verfügbar unter: https://regulaforensics.com/news/deepfake-fraud-costs/[Abruf: 16.07.2025].

Security Insider (2025): User and Entity Behavior Analytics (UEBA). Verfügbar unter: https://www.security-insider.de/was-ist-user-and-entity-behavior-analytics-ueba-a-983974/ [Abruf: 16.07.2025].

Surfshark (2025): Deepfake Statistics. Verfügbar unter: https://surfshark.com/research/study/deepfake-statistics [Abruf: 02.08.2025].

T3N (2025): Deepfakes in Echtzeit entlarven – Intel entwickelt FakeCatcher. Verfügbar unter: https://t3n.de/news/deepfakes-in-echtzeit-entlarven-intel-entwickelt-fakecatcher-1513086/ [Abruf: 02.08.2025].

Tehtris (2025): Deepfake-as-a-Service Threat Intelligence Report. Verfügbar unter: https://www.datensicherheit.de/wp-content/uploads/tehtris-deepfake-as-a-service-threat-intelligence-report.pdf [Abruf: 16.07.2025].

US Department of Homeland Security (2025): Impacts of Adversarial Use of Generative AI on Homeland Security. In: Preparedness Series, Januar 2025. Verfügbar unter: https://www.dhs.gov/archive/science-and- technology/publication/impacts-adversarial-use-generative-ai-homeland-security [Abruf: 02.08.2025].

cyberintelligence.institute
MesseTurm
Friedrich-Ebert-Anlage 49
D-60308 Frankfurt am Main

www.cyberintelligence.institute
info@cyberintelligence.institute

+49 69 505034602

Date of Publication: 10/2025