// KI-Hub VON STRUKTURAFLOW

IT-SECURITY

KI-gestützte Social-Engineering-Angriffe erkennen

Deepfake-Stimmen, LLM-Phishing, Persona-Synthese: Wie Mitarbeiter in österreichischen KMU KI-Angriffe erkennen und was NIS2/DSGVO verlangen.

AUTOR
Strukturaflow-Team

VERÖFFENTLICHT
28. Juni 2026

ONLINE LESEN
https://wissen.strukturaflow.it.com/ki-gestuetzter-social-engineering-angriff-mitarbeiter-erkennen-oesterreich/

Ein Mitarbeiter eines Wiener Handelsunternehmens erhält am Dienstagvormittag eine WhatsApp-Sprachnachricht. Die Stimme klingt exakt wie jene der Geschäftsführerin: dringend, vertraut, klar in der Anweisung — 14.000 € sollen noch heute auf ein neues Lieferantenkonto überwiesen werden. Er überweist. Die Stimme war synthetisch, erzeugt aus einem drei Sekunden langen Audioausschnitt eines YouTube-Interviews.

KI hat die Qualität von Social-Engineering-Angriffen so grundlegend verändert, dass klassische Warnsignale nicht mehr funktionieren. Schlechtes Deutsch, merkwürdige Absenderadressen, generische Ansprachen — das waren die verlässlichen Erkennungsmerkmale. KI-generierte Angriffe von 2024 und 2025 enthalten keinen dieser Fehler mehr.

Dieser Artikel erklärt, was KI-gestützte Angriffe technisch tun, woran Mitarbeiter sie trotzdem noch erkennen können, und welche konkreten Schritte österreichische KMU jetzt setzen sollten — inklusive NIS2-Rechtsrahmen und einem Simulationsszenario, das Sie heute im Team-Meeting durchführen können.

Was KI bei Social-Engineering-Angriffen konkret verändert

Klassisches Social Engineering arbeitete mit Masse: Millionen identische Phishing-Mails, von denen ein kleiner Prozentsatz ankam. Der Aufwand pro Opfer war minimal, die Qualität entsprechend erkennbar.

KI dreht dieses Verhältnis um. Der Aufwand pro Angriff bleibt gering — aber die Qualität steigt auf ein Niveau, das früher nur für hochspezialisierte, manuelle Angriffe auf Großkonzerne möglich war. Drei konkrete Techniken sind dafür verantwortlich:

LLM-generierte Spear-Phishing-Mails: Ein Large Language Model (KI-Sprachmodell wie GPT-4 oder ähnliche) kann aus öffentlich verfügbaren Daten — LinkedIn-Profil, Unternehmenswebsite, Presseaussendungen, Handelsregister — eine personalisierte E-Mail erstellen, die Projektname, Lieferantenname und Kollegenvorname enthält. Das klingt nicht wie Spam. Es klingt wie eine interne Mail.

Deepfake-Voice und Audio-Cloning: Moderne Voice-Cloning-Dienste benötigen drei bis fünf Sekunden Audiomaterial, um eine überzeugende Stimme zu replizieren. Quellen gibt es genug: WhatsApp-Sprachnotizen, Podcast-Auftritte, YouTube-Interviews, Messevideos. Das Ergebnis ist in Echtzeit generierbar — ein Anruf mit synthetischer Stimme ist heute technisch kein Hinderniss mehr.

Persona-Synthese: Gefälschte LinkedIn-Profile mit KI-generierten Profilfotos (erkennbar an bestimmten Mustern, aber auf den ersten Blick glaubwürdig), angereichert mit echten Unternehmensdetails aus öffentlichen Quellen. Ein Angreifer gibt sich als neuer Lieferant, Headhunter oder Behördenmitarbeiter aus — und hat eine vollständige digitale Identität, die eine flüchtige Prüfung übersteht.

Zur Einordnung: Das sind keine Science-Fiction-Szenarien. Die Werkzeuge dafür sind kostenfrei oder für wenige Euro monatlich verfügbar und erfordern kein technisches Vorwissen.

Die fünf häufigsten KI-Angriffsmuster auf KMU in Österreich

1. Spear-Phishing mit Firmenwissen

Die E-Mail kennt den laufenden Projektname, nennt den zuständigen Kollegen beim Vornamen und referenziert eine echte Lieferantenbeziehung. Sie bittet — mit leichtem Zeitdruck — darum, einen Anhang zu prüfen oder einer Zahlung zuzustimmen.

Erkennungsmerkmal: Eine unerwartete Handlungsaufforderung, die mit Zeitdruck kombiniert wird. Seriöse interne Prozesse erzeugen selten echten Sofortdruck per E-Mail. Prüfen Sie: Wurde diese Art von Anfrage je über genau diesen Kanal gestellt?

2. CEO-Fraud via KI-Stimme (Voice-Cloning)

Eine kurze Sprachnachricht oder ein Anruf von der Nummer der Führungskraft — oder einer ähnlich aussehenden Nummer — mit klarer Anweisung: Zahlung freigeben, Zugang gewähren, Vorgang diskret halten.

Wie Deepfake-Angriffe im Detail ablaufen und was Betroffene in Österreich tun können, beschreiben wir im Artikel Deepfake-Betrug bei österreichischen KMU.

Erkennungsmerkmal: Ungewöhnlicher Kanal kombiniert mit einer Bitte um sofortige Zahlung oder Zugangsdaten-Weitergabe. Die Stimme allein ist kein verlässlicher Beweis mehr — der Prozess ist es.

3. Gefälschte Lieferanten-Kommunikation

Eine täuschend echte E-Mail eines bestehenden Lieferanten teilt mit, dass sich die Bankverbindung geändert hat. Absenderdomain unterscheidet sich um einen Buchstaben vom echten Lieferanten (z. B. „lieferant-gmbh.com“ statt „lieferantgmbh.com“).

Erkennungsmerkmal: Jede Änderung von Zahlungsdaten oder Stammdaten, die über E-Mail kommt, muss telefonisch über die bekannte — nicht die in der E-Mail genannte — Nummer bestätigt werden. Ohne Ausnahme.

4. IT-Support-Impersonation (Helpdesk-Fraud)

Ein Anruf vom „internen IT-Support“ oder „Microsoft/Google-Support“ — oft zeitlich abgestimmt auf bekannte Störungen bei Cloud-Diensten. Der Anrufer bittet um einen MFA-Code, ein Passwort oder Remote-Zugriff auf den Rechner.

Erkennungsmerkmal: Kein legitimer IT-Dienst fragt per Anruf nach MFA-Codes oder Passwörtern. Niemals. Das ist die einfachste Regel, die gleichzeitig am häufigsten gebrochen wird.

5. Recruiting- und Bewerbungsangriff

Eine täuschend echte Bewerbungsmail an die HR-Verantwortliche enthält einen Anhang — oft als PDF getarnte .exe-Datei oder ein Word-Dokument mit aktiviertem Makro. Gerade kleinere Unternehmen mit öffentlich einsehbaren Stellenausschreibungen sind Ziel dieser Angriffe.

Erkennungsmerkmal: Dateitypen prüfen, bevor sie geöffnet werden. Makros in Office-Dokumenten von unbekannten Absendern grundsätzlich nicht aktivieren. Bewerbungsanhänge idealerweise in einer isolierten Umgebung öffnen.

Woran Mitarbeiter KI-Angriffe heute noch erkennen können

Die Qualität von KI-generierten Angriffen hat zugenommen — die psychologischen Manipulationsmuster dahinter haben sich nicht verändert. KI kann Sprachqualität und Personalisierung verbessern, aber sie muss trotzdem dieselben Hebel benutzen, die Social Engineering seit Jahrzehnten nutzt: Zeitdruck, Autoritätsgefälle, ungewöhnlicher Kanal, Geheimhaltungsappell.

Das ist der Ansatzpunkt für Mitarbeiter: nicht den Inhalt bewerten, sondern die Struktur der Anfrage.

Checkliste — Sechs Fragen vor jeder unerwarteten Anfrage:

- [ ] Kommt die Anfrage über einen ungewohnten Kanal (WhatsApp statt E-Mail, Anruf statt internes Ticket)?
- [ ] Wird Zeitdruck erzeugt („sofort“, „noch heute“, „bevor ich in die Besprechung muss“)?
- [ ] Werde ich gebeten, den Vorgang geheim zu halten oder nicht mit Kollegen zu sprechen?
- [ ] Stimmt die Absenderadresse vollständig überein — nicht nur der Anzeigename, sondern die tatsächliche Domain?
- [ ] Wurde eine Zahlung oder Zugangsdaten-Herausgabe über genau diesen Kanal vorher je gemacht?
- [ ] Kann ich die Person über einen zweiten, unabhängigen Kanal verifizieren (Rückruf auf bekannte Nummer)?

Besonders wichtig beim Thema Deepfake-Voice: Die Qualität synthetischer Stimmen hat ein Niveau erreicht, bei dem Sinneseindrücke allein kein verlässliches Erkennungsmittel mehr sind. Die Antwort darauf ist kein besseres Hinhören — sondern ein Prozess, der unabhängig von der Überzeugungskraft einer Stimme gilt. Wenn der Prozess sagt „Rückruf vor Freigabe“, gilt das auch dann, wenn die Stimme täuschend echt klingt.

Das 5-Minuten-Simulationsszenario für Ihr nächstes Team-Meeting

Externe Trainer und spezialisierte Tools sind hilfreich — aber nicht Voraussetzung für eine erste Übung. Diese zwei Szenarien können Sie heute in jedem Team-Meeting durchführen, ohne Budget und ohne technisches Wissen.

Szenario A — Spear-Phishing per E-Mail:

Die Geschäftsführung schickt eine interne Test-Mail an einen Mitarbeiter. Inhalt: „Ich bin gerade im Gespräch und brauche dringend, dass du die angehängte Rechnung von [echter Lieferantenname] bis 14 Uhr freigibst — Matthias ist informiert.“ Die Mail kommt von einer leicht abweichenden Absenderadresse (z. B. privates Konto). Kein Anhang nötig — die Reaktion des Mitarbeiters ist das Ziel, nicht der Klick.

Szenario B — Helpdesk-Impersonation per Anruf:

Eine Kollegin ruft einen anderen Mitarbeiter von einer unbekannten Nummer an und gibt sich als „IT-Support“ aus. Sie erklärt, es gebe ein Problem mit dem Microsoft-365-Konto und bittet um den aktuellen MFA-Code zur Verifikation.

Nachbesprechung (3 Fragen): 1. Was genau hat Sie an der Anfrage zögern lassen — oder eben nicht? 2. Welchen Schritt hätten Sie als nächstes unternommen? 3. Über welchen Kanal hätten Sie die Anfrage verifiziert?

Wichtig: Diese Übung hat keinen Sanktionscharakter. Wer darauf hereinfällt, hat nichts falsch gemacht — das System hat versagt, nicht die Person. Ziel ist eine offene Diskussion über Erkennungsmerkmale und interne Prozesse.

Für professionelle Simulationskampagnen mit messbaren Ergebnissen stehen Tools wie KnowBe4 oder Proofpoint Security Awareness zur Verfügung, die auch in österreichischen KMU eingesetzt werden. GoPhish ist eine kostenfreie Open-Source-Alternative für kleinste Teams.

Österreichischer Rechtsrahmen: Was NIS2 und DSGVO von Ihnen verlangen

Mit dem Netz- und Informationssystemsicherheitsgesetz (NISG 2024) hat Österreich die EU-NIS2-Richtlinie in nationales Recht umgesetzt. Für KMU sind zwei Schwellenwerte relevant: Ab 50 Mitarbeitern oder 10 Millionen Euro Jahresumsatz gilt ein Unternehmen als „wichtige Einrichtung" und fällt direkt unter die Anforderungen — darunter technische und organisatorische Sicherheitsmaßnahmen, zu denen Mitarbeiterschulungen implizit gehören.

Einen ausführlichen Überblick, welche Pflichten konkret entstehen und was auch für kleinere Unternehmen relevant ist, bietet unser Artikel NIS2 / NISG 2026: Auch ohne direkte Pflicht relevant für KMU?.

Für Kleinstunternehmen unter diesen Schwellenwerten gilt: Eine direkte NIS2-Pflicht besteht i.d.R. nicht. Aber als Zulieferer einer betroffenen Einrichtung können Sicherheitsanforderungen über Vertrag auferlegt werden (Supply-Chain-Sicherheit). Die enthaltenen Maßnahmen — Vier-Augen-Prinzip, Schulungspflicht, Notfallprozesse — schaffen unabhängig davon Rechtssicherheit und sind schlicht sinnvoll.

DSGVO-Meldepflicht: Ein erfolgreicher Social-Engineering-Angriff ist in vielen Fällen eine Datenpanne im Sinne der DSGVO — etwa wenn Zugangsdaten kompromittiert wurden oder personenbezogene Daten abgeflossen sind. Nach aktueller Rechtslage besteht in diesem Fall eine Meldepflicht an die österreichische Datenschutzbehörde (DSB) binnen 72 Stunden. Die Meldung online einzureichen ist über dsb.gv.at möglich.

Haftung der Geschäftsführung: Bei grober Fahrlässigkeit — etwa wenn trotz bekannter Bedrohungslage keinerlei Schutzmaßnahmen implementiert wurden — können nach aktuellem Zivilrecht Haftungsansprüche entstehen. Das ist kein theoretisches Risiko mehr, sobald Schadenssummen wie im Eingangsbeispiel im Raum stehen.

Weiterführende Informationen und Orientierungshilfen bieten RTR.gv.at, der WKO-Sicherheitsleitfaden sowie dsb.gv.at — alle kostenlos zugänglich.

Sofortmaßnahmen für KMU ohne eigene IT-Abteilung

Diese fünf Maßnahmen können Sie ohne externes Budget, ohne IT-Abteilung und ohne Wochen der Vorbereitung umsetzen:

1. Vier-Augen-Prinzip für Zahlungen und Stammdatenänderungen: Jede Zahlung ab einem definierten Betrag (z. B. 2.000 €) und jede Änderung von Bankverbindungen oder Kontaktdaten erfordert eine zweite Bestätigung — persönlich oder telefonisch über eine verifizierte Nummer. Diesen Prozess schriftlich festhalten.

2. Rückruf-Policy: Bei verdächtigen Anfragen wird immer über eine bekannte, verifizierte Nummer zurückgerufen — niemals über die in der verdächtigen Nachricht genannte. Diese Regel gilt auch dann, wenn die Anfrage von einer vertrauten Person zu kommen scheint.

3. MFA als Pflicht, nicht als Option: Mehrstufige Authentifizierung für E-Mail-Zugänge und Cloud-Dienste ist keine Empfehlung, sondern Grundschutz. Ob Ihr aktuelles Passwort-Management darüber hinaus ausreicht, lesen Sie in unserem Artikel Passwort-Sicherheit: Reicht der Chrome-Browser-Check wirklich?.

4. Minimalschulung — mindestens einmal jährlich: Eine einmalige Schulung ist besser als keine — reicht aber langfristig nicht. Empfehlung: halbjährliche Auffrischung, zusätzlich nach neuen Angriffswellen oder internen Vorfällen. Die Team-Meeting-Simulation aus dem vorigen Abschnitt kann diese Funktion übernehmen.

5. Notfallkontakt definieren: Wer wird bei Verdacht als Erstes informiert? Diese Frage muss beantwortet sein, bevor ein Angriff passiert — nicht danach. Einen internen Ansprechpartner und einen externen IT-Dienstleister/MSP als Eskalationspfad festlegen und schriftlich kommunizieren.

Externe Dienstleister und Managed Service Provider können darüber hinaus simulierte Phishing-Kampagnen und Schulungsprogramme übernehmen — gerade wenn intern die Zeit oder Expertise fehlt, diese kontinuierlich durchzuführen.

Häufige Fragen

Kann KI-generiertes Phishing von normalen Spam-Filtern erkannt werden?

Zuverlässig nicht. LLM-generierte Mails enthalten keine klassischen Spam-Muster — keine verdächtigen Absender-Domains, keine Tippfehler, keine generischen Ansprachen. Verhaltensbasierte Erkennungssysteme, die auf Anomalien im Absenderverhalten oder ungewöhnlichen Weiterleitungsregeln reagieren, sind verlässlicher. Sie erfordern aber entsprechende Tools und Konfiguration — ein Standard-Spam-Filter reicht nicht.

Was tue ich, wenn ein Mitarbeiter bereits auf einen Angriff hereingefallen ist?

Sofortmaßnahmen in dieser Reihenfolge: Zugänge sperren, IT-Dienstleister informieren, Sachverhalt so genau wie möglich dokumentieren. Bei Datenverlust oder kompromittierten personenbezogenen Daten: DSGVO-Meldepflicht an die österreichische Datenschutzbehörde (DSB) binnen 72 Stunden prüfen. Keine Bestrafung des betroffenen Mitarbeiters — Angriffe dieser Qualität treffen auch erfahrene Personen.

Gilt NIS2 auch für mein 12-Personen-Unternehmen?

Direkt verpflichtend in der Regel nicht (Schwellenwert: 50 Mitarbeiter oder 10 Mio. € Umsatz). Indirekt können Sie als Zulieferer einer betroffenen Einrichtung über Verträge dennoch Anforderungen auferlegt bekommen. Die Maßnahmen des NISG 2024 sind unabhängig davon empfehlenswert.

Wie teuer sind professionelle Phishing-Simulationen für KMU?

Einstiegslösungen wie KnowBe4 Starter oder Proofpoint Essentials liegen bei ca. 10–25 € pro Nutzer und Jahr. Für kleinste Teams gibt es mit GoPhish eine kostenfreie Open-Source-Alternative.

Reicht einmal jährliche Schulung aus?

Nein. Mindestens halbjährliche Auffrischung wird empfohlen, zusätzlich anlassbezogen nach neuen Angriffswellen oder internen Vorfällen. KI-gestützte Angriffstechniken entwickeln sich schnell — Schulungsinhalte müssen mithalten.

Nächste Schritte

Viele der in diesem Artikel beschriebenen Maßnahmen sind bekannt — sie scheitern in der Praxis aber häufig an einer konkreten Lücke: fehlende Zeit zur Priorisierung, Unsicherheit wo anfangen, oder kein klares Bild davon, wie der eigene Reifegrad im Vergleich zur tatsächlichen Bedrohungslage aussieht.

Ein einziger erfolgreicher Angriff kostet in den meisten Fällen mehr als ein Jahr präventiver Maßnahmen — finanziell, aber auch in Bezug auf Vertrauen und Reputationsschaden.

Im kostenlosen Beratungsgespräch mit dem Strukturaflow-Team analysieren wir gemeinsam Ihre aktuelle Situation: Welche Prozesse sind bereits vorhanden, wo gibt es die drängendsten Lücken, und welche zwei oder drei Maßnahmen würden sofort den größten Schutz bringen. Das Gespräch dauert 30 Minuten und erfordert kein Vorwissen — und kein Folgemandat. Es geht darum, Ihnen Orientierung zu geben, nicht um einen Vertragsabschluss.

NÄCHSTER SCHRITT

Mehr praktische KI-Anleitungen für KMU

Dieser Artikel ist Teil des KI-Hubs von Strukturaflow — einer deutschsprachigen Plattform für den praktischen KI-Einsatz in kleinen und mittleren Unternehmen.

https://wissen.strukturaflow.it.com