SGO VEREIN

DIE SGO IST DIE COMMUNITY
FÜR ORGANISATION UND MANAGEMENT

www.sgo-verein.ch

SGO TALK I 25.06.2026

«KI & CYBERSECURITY – FLUCH UND SEGEN»

Fabian Wipf I Chief Information Security Officer / CISO, LAKE Solutions AG
Viktor Dötzel I Senior Security Engineer, LAKE Solutions AG

www.sgo-verein.ch

ÜBER DIE SGO

Die Community für Organisation & Management

Interdisziplinär und branchenübergreifend bringt die SGO Fach- und Führungspersonen zusammen, die Organisation und Management aktiv mitgestalten.

«Lerne aus Erfahrung.»

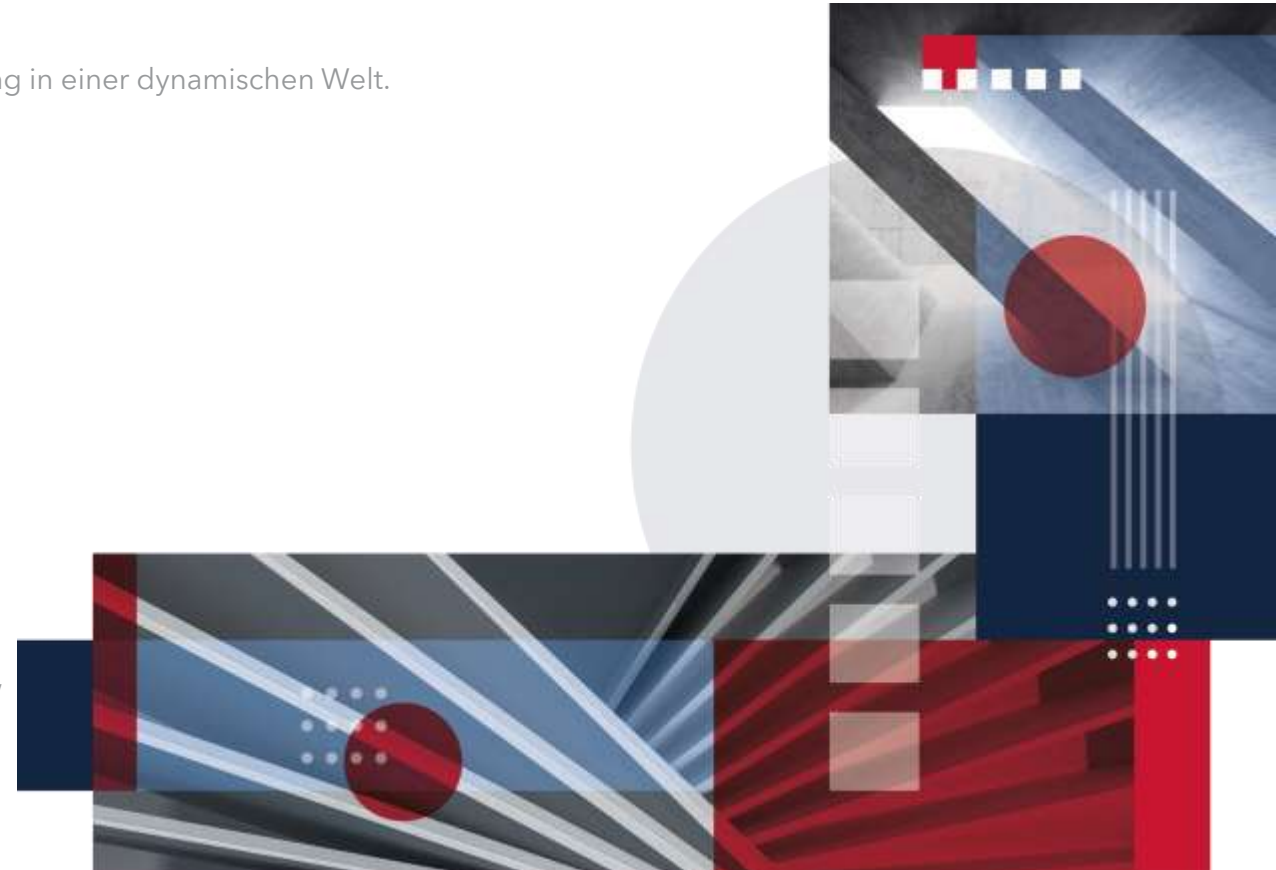
Wir schaffen Raum für Austausch auf Augenhöhe, neue Perspektiven und Orientierung in einer dynamischen Welt.

Was die SGO ausmacht:

- praxisnaher Erfahrungsaustausch
- relevante Themen aus der Unternehmenspraxis
- Wirkung in den Schnittmengen von Strategie, Struktur und Kultur
- **Vernetzung von Fach- und Führungspersonen über Branchen hinweg**
- Organisation ganzheitlich denken – in den Schnittmengen von Strategie, Struktur und Kultur

rund 25 Veranstaltungen pro Jahr

mit Einblicken in Unternehmen, konkreten Praxisbeispielen und aktuellem Know-how



LAKE
A RICOH Company

Fabian Wipf
Chief Information
Security Officer
fabian.wipf@lake.ch

Viktor Dötzel
Senior Security Engineer
viktor.doetzel@lake.ch

RICOH | Cloud & Cyber

Agenda

LAKE
A RICOH Company

01

Teil 1 — KI als Werkzeug der Angreifer

Deepfakes, CEO-Fraud & die neue Bedrohungslage

JETZT

02

Teil 2 — Die Antwort der Organisation

Governance, Policies & Defense

Welches Szenario bereitet euch am meisten Kopfschmerzen?

AI Agents

Das ist Summer Yue.

Direktorin für AI Alignment
bei Meta Superintelligence Labs

Ihr Job: KI sicher machen.

Videokonferenz „Dringende Finanzfreigabe“

KI · FAKE „CFO“

KI · FAKE „Kollege“

KI · FAKE „Kollegin“

KI · FAKE „Senior Manager“

KI · FAKE „Finanzleitung“

ECHT Arup-Mitarbeiter

KI · FAKE „Kollege“

KI · FAKE „Kollegin“

Alle Teilnehmer ausser dem Mitarbeiter waren KI-Deepfakes.

Ein Anruf, der 25 Millionen kostete

25 Mio.

US-Dollar — in 15 Überweisungen

- Eine komplette Videokonferenz mit „Kollegen" und dem „CFO"
- **Jede einzelne Person im Call war ein KI-Deepfake**
- Aufgeflogen erst Tage später beim Rückruf in der Zentrale

„Der Mitarbeiter hatte zunächst einen Betrug vermutet — der überzeugende Live-Videocall hat seine Skepsis vollständig ausgeräumt."

Ingenieurbüro Arup, Hongkong, 2024 — bis heute einer der bekanntesten Deepfake-Betrugsfälle.

Quelle: World Economic Forum / Institute for Financial Integrity, 2024–2025

Was sich verändert hat

Nicht „mehr vom Gleichen" — ein neues Niveau bei Tempo, Menge und Glaubwürdigkeit.

5 Min.

für eine überzeugende Phishing-Mail per KI — früher einige Stunden Handarbeit.

4% → 56%

Sprung KI-generierter Phishing-Angriffe innerhalb weniger Wochen (Dez. 2025).

kaum Fehler

Tippfehler & holprige Anrede als Warnsignal sind praktisch verschwunden.

Quellen: IBM X-Force; Hoxhunt Phishing Trends 2025/2026

Phishing: früher vs. heute

Zwei Mails, die das Gleiche wollen — Ihr Geld. Welche würde Sie erwischen?

FRÜHER

Von: security@paypal-verify-konto.com

Betreff: DRINGEND!!! Ihr Konto wurde gesperrt

Sehr geehrte Kunde,
Ihres Konto wurde wegen verdächtige Aktivität gespert.
Klicken Sie sofort hier um zu verifizieren Ihre Daten, sonst wird gelöscht in 24 Stunde.

Red-Flags: schlechtes Deutsch, generische Anrede, auffällige Absender-Adresse, Drohung & Zeitdruck

HEUTE — KI-generiert

Von: t.berger@firmenname-finance.ch

Betreff: Re: Freigabe Zahlung Projekt Hafencity

Hallo Frau Klein, wie eben kurz besprochen — können Sie die Zahlung an den Lieferanten für die Hafencity-Phase noch Heute anstossen? Thomas ist bis 15 Uhr im Termin, Details finden Sie im Anhang.
Danke!
Viele Grüsse, T. Berger

Problem: korrekte Anrede, echter Projektbezug, sehr gutes Deutsch, plausibler Kollege — kein einziges klassisches Warnsignal.

Deepfakes: erschreckend wenig nötig

20–30 Sekunden

Audio reichen, um eine Stimme überzeugend zu klonen.

Und dieses Material ist längst öffentlich:

- **Interviews & Podcasts**
- **Firmen- & Messevideos**
- **Webinare / Online-Calls**
- **LinkedIn- & Social-Clips**

Für eine Führungskraft existieren oft Stunden an Material. Quelle: FBI / Brightside AI 2025–2026

Spot the Deepfake

LAKE
A RICOH Company

Warum das funktioniert

Der Angriff zielt nicht auf Unwissen — sondern auf menschliche Reflexe.

Autorität

„Die Chefin sagt es."
Wir hinterfragen Vorgesetzte ungern.

Zeitdruck

„Sofort, sonst …"
Eile schaltet das kritische Denken aus.

Vertrauen ins Sehen

„Ich sehe & höre die Person doch."
Diese Annahme ist gekippt.

Deshalb hilft reines „Augen auf"-Training nur begrenzt — der Reflex schlägt das Wissen.

Der Werkzeugkasten der Angreifer

Was früher Spezialwissen brauchte, übernimmt heute zunehmend die KI — schneller und günstiger.

Angriff wird Massenware

- KI sucht selbstständig nach Schwachstellen
- „Tester"-Systeme schlagen erstmals viele menschliche Profis
- Wenig Können nötig — Angriffe werden mietbar
- Aus „wenige Top-Hacker" wird „sehr viele"

Beweise unter Verdacht

- Können wir Video & Audio noch als Beweis trauen?
- „Sehen ist glauben" gilt nicht mehr blind
- Echtheits-Prüfung von Aufnahmen wird zur Kerndisziplin
- Spurensuche braucht jetzt selbst KI-Werkzeuge

GenAI-Apps und -Agenten schaffen Sicherheitsrisiken

LAKE
A RICOH Company

Wildwuchs von Agenten & Ressourcenzugriff

82 %

der Führungskräfte planen, in den nächsten 12–18 Monaten Agenten einzusetzen, um den Bedarf an Arbeitskräftekapazität zu decken[3]

Datenlecks & übermässiges Teilen

80 %

der Führungskräfte nannten das Abfliessen sensibler Daten als ihre grösste Sorge[1]

Neue KI-Bedrohungen & Schwachstellen

88 %

der Organisationen sind besorgt über indirekte Prompt-Injection-Angriffe[2]

Einhaltung von Vorschriften

55 %

der Führungskräfte fehlt das Verständnis dafür, wie KI reguliert wird und werden soll, und sie suchen nach Orientierung[1]

1. First Annual Generative AI study: Business Rewards vs. Security Risks, Q3 2023, ISMG, N=400
2. How to Secure Custom-Built AI Agents, Gartner, 17. März 2025, Dionisio Zumerle, Jeremy D'Hoinne
3. Microsoft Work Trend Index Survey 2025

Angriff über die KI-Zusammenfassung

Beispiel: "ChatGPhish" – eine modifizierte Webseite genügt

1

Versteckter Befehl

Angreifer baut eine unsichtbare Anweisung in eine Webseite ein.

2

KI fasst zusammen

Mitarbeiter lässt die Seite ganz normal von der KI zusammenfassen.

3

Vertrauter Köder

Die KI zeigt Link & QR-Code als seriös in ihrer Oberfläche an.

4

Klick / Scan

Opfer wird auf Phishing geleitet — Zugangsdaten weg.

Kein Anhang. Keine verdächtige Mail.

Nur eine ganz normale KI-Zusammenfassung — das Vertrauen in die KI ist die eigentliche Schwachstelle.

Quelle: Permiso Security / The Hacker News, Mai 2026

Pause

Agenda

LAKE
A RICOH Company

01

Teil 1 — KI als Werkzeug der Angreifer

Deepfakes, CEO-Fraud & die neue Bedrohungslage

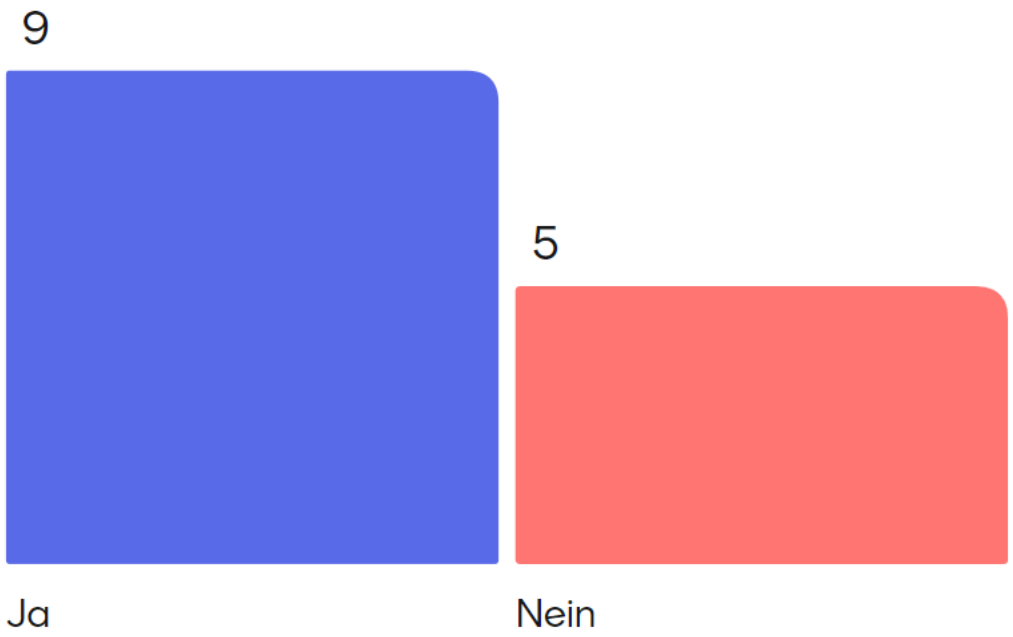
02

Teil 2 — Die Antwort der Organisation

Governance, Policies & Defense

JETZT

Wart ihr schon direkt/indirekt betroffen?



Dieselbe Technologie — jetzt auf unserer Seite

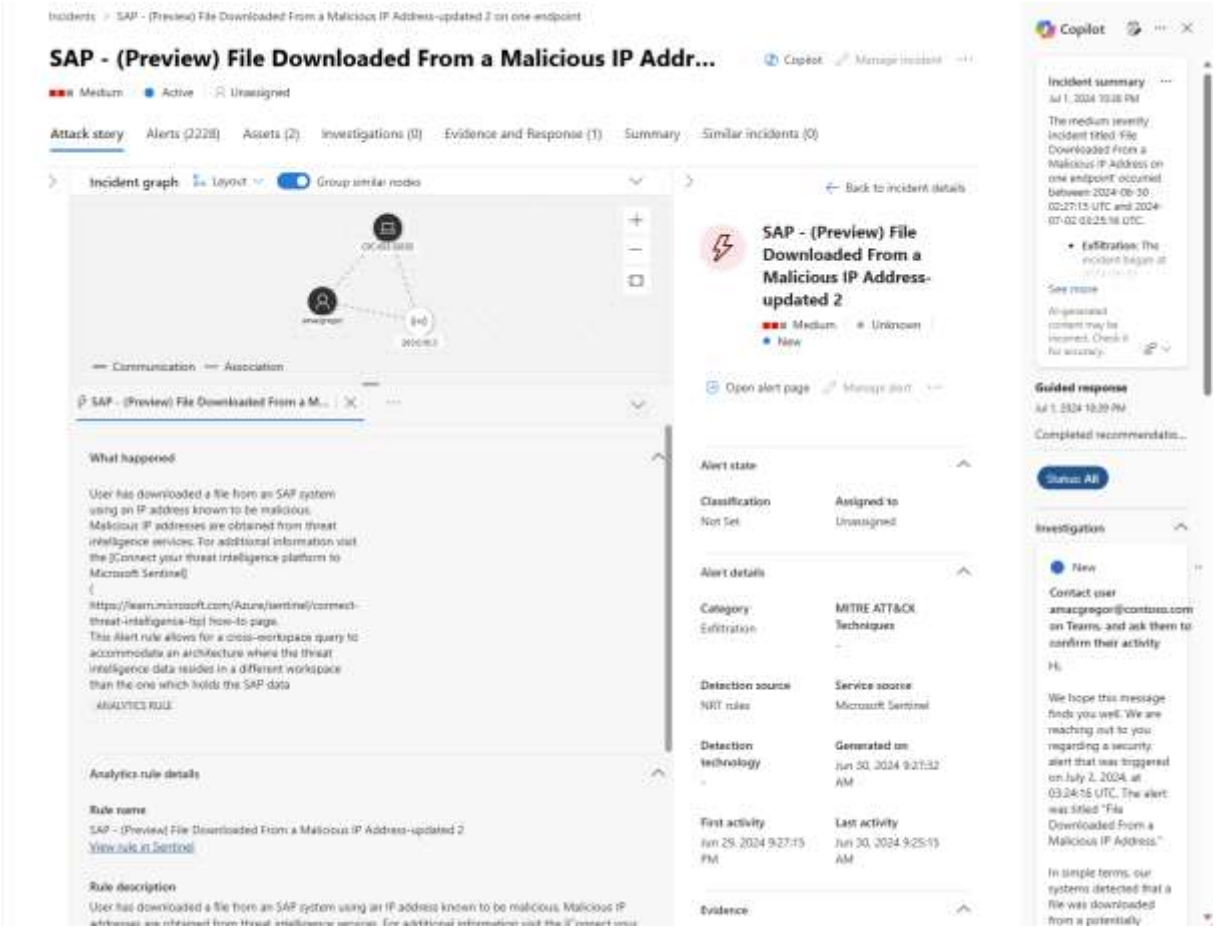
In Block 1 war KI die Waffe. Jetzt drehen wir sie um: KI als Verteidigungswerkzeug.

KI-gestützte Bedrohungserkennung

LAKE
A RICOH Company

Das Sicherheitsteam bekommt sein eigenes KI-Werkzeug.

- SOCs nutzen KI, um Anomalien in Echtzeit zu erkennen.
- Millionen Events täglich — kein Mensch wertet das manuell aus
- Mustererkennung statt starrer Signaturen — auch unbekannte Angriffe.
- Beispiele: Microsoft Sentinel, CrowdStrike Falcon.



Warum der Mensch allein nicht mehr reicht

Das gleiche Datenvolumen — zwei Wege, damit umzugehen.

MANUELL

- Analyst:in sichtet Alarme einzeln
- Bekannte Signaturen, feste Schwellen
- Reaktion in Stunden bis Tagen
- Alert-Müdigkeit bei Tausenden Meldungen

KI-GESTÜTZT

- Automatische Priorisierung & Korrelation
- Verhaltensbasiert — erkennt Abweichungen
- Reaktion in Sekunden bis Minuten
- Mensch entscheidet, KI filtert das Rauschen

Awareness: vom Pflicht-Kurs zum Micro-Learning

Ein Kreislauf, der sich an reales Verhalten anpasst — statt einmal jährlich für alle.

1. Simulieren

Realistische, KI-generierte Phishing-Tests, individuell je Person.

2. Messen

Wer klickt, wer meldet? Verhalten anonymisiert auswerten.

3. Anpassen

Lernpfad pro Person: mehr Übung, wo die Schwachstelle liegt.

4. Wiederholen

Kontinuierlich statt Einmal-Schulung.

Wirkung: aus dem „schwächsten Glied" wird ein aktiver Sensor — Mitarbeitende melden Verdächtiges.

Von der Erkennung zur Reaktion

KI beschleunigt die Reaktion — der Mensch bleibt in der Entscheidung.

Triage & Priorisierung

KI sortiert Tausende Alarme, hebt das wirklich Kritische hervor.

Kontext anreichern

Zusammenhänge über Systeme hinweg automatisch verknüpfen.

Erste Eindämmung

Vorschläge oder automatisierte Sofortmassnahmen, z.B. Konto sperren.

Mensch entscheidet

Analyst:in prüft & eskaliert — KI übernimmt die Fleissarbeit.

Technik allein reicht nicht

Was uns wirklich schützt, sind klare Weisungen — die Governance hinter der Technik.

Fünf Bausteine guter KI-Governance

LAKE
A RICOH Company

Weisungen, Massnahmen und Pflichten — gleich vertiefen wir jeden einzeln.

AI Usage Policy	Welche Daten dürfen in welche KI-Tools?	Weisung
Deepfake-Protokolle	Rückrufpflicht bei ungewöhnlichen Anweisungen.	Massnahme
Shadow AI	Nicht-genehmigte KI-Tools inventarisieren.	Massnahme
Regulatorik	revDSG / DSG, FINMA, EU AI Act & NIS2.	Pflicht
Incident Response	KI-Szenarien in Notfallpläne aufnehmen.	Massnahme

AI Usage Policy

Klare Regeln, welche Daten in welche KI-Tools eingegeben werden dürfen.

- Datenklassen definieren (öffentlich / intern / vertraulich).
- Pro Klasse festlegen: welches Tool ist erlaubt?
- Unterschied beachten: öffentliche Tools vs. Enterprise-Tenant.
- Verbindlich machen — als Weisung, nicht als Empfehlung.

PRAXIS-TIPP

Eine konkrete Klausel zeigen wirkt mehr als die abstrakte Regel: „Vertrauliche Kundendaten nur in den Firmen-Tenant von Tool X."

Deepfake-Protokolle

LAKE
A RICOH Company

Rückrufpflicht bei ungewöhnlichen Anweisungen per Video oder Audio.

- Zweiter Kanal: bei Zahlungs- oder Datenanfragen immer zurückrufen.
- Bekannte Nummer nutzen — nie die aus der Nachricht.
- Codewörter für besonders sensible Freigaben etablieren.
- Mitarbeitende ermutigen, Anweisungen zu hinterfragen.

PRAXIS-TIPP

Die wirksamste Massnahme ist die simpelste: ein zweiter, bekannter Kanal. Kein Deepfake übersteht einen Rückruf auf die hinterlegte Nummer.

Shadow AI

LAKE
A RICOH Company

Inventarisierung nicht-genehmigter KI-Tools im Unternehmen.

- Überblick verschaffen: welche Tools sind im Einsatz?
- Risiko: Firmendaten landen unkontrolliert in externen Diensten.
- Lieber freigegebene Alternative anbieten als pauschal verbieten.
- Regelmässig neu erheben — die Tool-Landschaft ändert sich schnell.

PRAXIS-TIPP

Frage ins Publikum: „Wissen Sie, in wie viele KI-Tools Ihre Mitarbeitenden gerade Firmendaten eingeben?"

Regulatorik im KI-Kontext

Schweizer und europäische Anforderungen, die KI-Nutzung betreffen.

- revDSG / DSG: Datenschutz bei KI-Verarbeitung in der Schweiz.
- FINMA: Outsourcing- und Auslagerungs-Anforderungen.
- EU AI Act: risikobasierte Pflichten je nach Anwendung.
- NIS2: erweiterte Cybersecurity-Pflichten in der EU.

PRAXIS-TIPP

Für ein CH-Publikum: revDSG zuerst, dann der Blick nach Europa. Nicht alles trifft jeden — Betroffenheit prüfen.

Incident Response

KI-spezifische Szenarien gehören in bestehende Notfallpläne.

- Deepfake-Vorfall: wer wird wie alarmiert?
- Datenabfluss über ein KI-Tool als eigenes Szenario.
- Verantwortlichkeiten und Eskalationswege definieren.
- Im Tabletop-Übungsplan mit durchspielen.

PRAXIS-TIPP

KI erfindet die Incident Response nicht neu — sie ergänzt vorhandene Pläne um neue Szenarien.

Eine 80%-Lösung heute

ist besser als eine 100%-Lösung in einem halben Jahr.

Klein anfangen

Mit einem Handlungsfeld starten, nicht mit allen fünf gleichzeitig.

Schritt für Schritt

Lieber etwas umsetzen und verbessern als auf das Perfekte warten.

Wir unterstützen gerne

Sie müssen das nicht allein machen — wir begleiten Sie dabei.

30-Tage-Roadmap — sofort starten

Konkret, machbar, ohne grosses Budget.

1 Woche 1

Sichtbarkeit

- Shadow-AI-Inventar starten
- Aktuelle Tool-Nutzung erheben

2 Woche 2–3

Regeln setzen

- AI Usage Policy entwerfen
- Deepfake-Rückrufregel definieren

3 Woche 4

Verankern

- Policy kommunizieren & schulen
- IR-Plan um KI-Szenarien ergänzen

KERNBOTSCHAFT

KI ist kein Silver Bullet —

aber wer KI nicht zur Verteidigung einsetzt, kämpft mit veralteten Waffen.

3 Take-aways

LAKE
A RICOH Company

Konkret und ab dieser Woche umsetzbar

1

Deepfake-Rückrufregel einführen

Bei ungewöhnlichen Audio-/Video-Anweisungen: zurückrufen, bevor freigegeben wird.

2

Shadow-AI-Inventar starten

Überblick verschaffen: in welche KI-Tools fliessen heute Firmendaten?

3

AI Usage Policy verankern

Als verbindliches Pflichtdokument — welche Datenklassen in welche Tools.

VIELEN DANK AN UNSERE COMMUNITY REFERIERENDEN UND PARTNER

Die SGO Community lebt vom Austausch unterschiedlichster Perspektiven – danke an unsere Community und Partner, die dies mit ihrer Unterstützung ermöglichen.

Berner Fachhochschule

LAKE
A RICOH Company

SGO BUSINESS SCHOOL

UNSERE NÄCHSTEN EVENTS

03.09.26
17:00 Uhr

TECHNOLOGIE & INFRASTRUKTUR

PROZESSINTELLIGENZ FÜR UNTERNEHMEN – AUTOMATISIERUNG KONKRET UMGESETZT

Veranstaltungsort folgt!

Event ansehen →

20.10.26
17:00 Uhr

MENSCHEN & KOMPETENZEN

DANK MEETINGKULTUR MOTIVIERTE, PRODUKTIVE UND LEISTUNGSSTARKE MITARBEITENDE UND ORGANISATION

SGO Practice Event

memox YOND | Albisriederstrasse 199, 8047 Zürich

Event ansehen →

27.10.26
17:00 Uhr

COMMUNITY EVENT

8. SGO COMMUNITY EVENT

Save the Date

Universität Bern | Hauptgebäude H4, Hochschulstrasse 4, 3012 Bern

Event ansehen →

www.sgo-verein.ch/events

BESTEN DANK

www.sgo-verein.ch

SGO Schweizerische Gesellschaft
für Organisation und Management

8000 Zürich

Tel: 044 809 99 55

welcome@sgo-verein.ch

www.sgo-verein.ch