CHANGING PERSPECTIVE
Austrian IT & CONSULTANTS DAY 2025
BLICKWINKEL ZUKUNFT

# CHANGING PERSPECTIVE

Austrian IT & CONSULTANTS DAY 2025

BLICKWINKEL ZUKUNFT

CHANGING PERSPECTIVE
Austrian IT & CONSULTANTS DAY 2025
BLICKWINKEL ZUKUNFT

# Vertrauen in KI

# Haben Sie Vertrauen in KI?

CHANGING PERSPECTIVE
BLICKWINKEL ZUKUNFT
Austrian IT & CONSULTANTS DAY 2025

# Haben Sie Vertrauen in KI?

Nur wer die Risiken kennt, kann Vertrauen schaffen

# Tim Hoffmann

- Studium der Wirtschaftswissenschaften mit Schwerpunkten in
  - Organisation
  - Informationsmanagement
  - Datenschutz-Controlling
- seit 2002 als Berater mit den Schwerpunkten Datenschutz und Informationssicherheit in KMU
- Coach für KI-Kompetenz

# Wer wir sind

## *Datenschutz*

Externe Datenschutzbeauftragung, Audits, E-Learning, Dienstleister-Auditierung etc.

## *Informationssicherheit*

Aufbau eines ISMS bis zur Zertifizierung, Informationssicherheitsbeauftragter etc.

## *Organisation / Strategie*

Beratungsleistungen bei konzeptionellen und strategischen Fragestellungen
(inkl. Hinweisgeberschutz oder KI-Kompetenz)

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen (Kurz-Überblick)

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen (Kurz-Überblick)

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

# Künstliche Intelligenz

**KI oder Marketing-Label?**
Die Grenzen sind fließend.

KI kann aus Daten lernen und Muster erkennen

bisherige Algorithmen basierte Systeme arbeiten mit festgelegten Regeln

Unterscheidung:

- Generative KI: erzeugt originelle Daten (Text, Bilder oder Musik)
- Diskriminative KI: erkennt Unterschiede zwischen verschiedenen Daten

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

**Rechtliche Anforderungen (Kurz-Überblick)**

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

# KI-Verordnung

*Der AI Act (KI-Verordnung) soll die Einführung von **menschen-zentrierten** und **vertrauenswürdigen** KI-Systemen fördern und gleichzeitig ein **hohes Maß an Schutz** für Gesundheit, Sicherheit und Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Schutz der Umwelt gewährleisten.*

Risikobasierter Ansatz

Ersetzt nicht die DSGVO

Ziel ist es, Vertrauen zu schaffen

# Akteure in KI-Verordnung

Anbieter („Provider“)

Produkthersteller („Product Manufacturer“)

Bevollmächtigter („Authorised Representative“)

Einführer („Importer“)

Händler („Distributor“)

Betreiber („Deployer“)

# Akteure in KI-Verordnung

**Anbieter („Provider“)** — „Hersteller“

Produkthersteller („Product Manufacturer“)

Bevollmächtigter („Authorised Representative“)

Einführer („Importer“)

Händler („Distributor“)

**Betreiber („Deployer“)** — „Nutzer“

**Vorsicht bei Customizing:**
Ab wann wird Betreiber zum Anwender?

# Risikostufen von KI-Systemen

| | | |
|---|---|---|
| inakzeptabel | **Verboten**, weil sie im Widerspruch zu den Werten der EU stehen | Social Scoring, unterschwellige Technologien zur (negativen) Beeinflussung, Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz |
| Hoch | Anforderungen an Anbieter und Betreiber | Personalmanagement, Kreditwürdigkeit, Vertragsabschlüsse |
| begrenzt | Risikomanagement durch Transparenzpflichten | Chatbots / generative KI für Bilder, Videos, Stimmen etc. |
| minimal | Freiwillige Verhaltens-Kodize | Spamfilter, Übersetzung, Rechtschreib- und Grammatikkorrekturen |

# Verpflichtungen für *Betreiber*

| **Beachte: keine vollständige Auflistung** | Hoch | begrenzt | minimal |
|---|---|---|---|
| KI-Kompetenz | ✓ | ✓ | ✓ |
| Transparenz gegenüber nachgelagerten Akteuren | ✓ | ✓ | |
| Verwendung des KI-Systems laut Betriebsanleitung | ✓ | | |
| Menschliche Aufsicht | ✓ | | |
| Überwachung des KI-Systems | ✓ | | |
| Meldung von schwerwiegenden Vorfällen | ✓ | | |
| Aufbewahrung von erzeugten Protokollen | ✓ | | |
| Sofern relevant: Datenschutz-Folgenabschätzung | ✓ | | |
| Zusammenarbeit mit zuständigen nationalen Behörden | ✓ | | |
| Recht auf Erläuterung der Entscheidungsfindung im Einzelfall | ✓ | | |
| Informationspflichten gegenüber der Arbeitnehmerinnen-Vertretung<br>*sofern Arbeitgeberin Hochrisiko-KI-Systeme am Arbeitsplatz einsetzt* | ✓ | | |

# Verpflichtungen für *Betreiber*

**Beachte: keine vollständige Auflistung**

| | Hoch | begrenzt | minimal |
|---|---|---|---|
| **KI-Kompetenz** | ✓ | ✓ | ✓ |
| Transparenz gegenüber nachgelagerten Akteuren | ✓ | ✓ | |
| Verwendung des KI-Systems laut Betriebsanleitung | ✓ | | |
| Menschliche Aufsicht | ✓ | | |
| Überwachung des KI-Systems | ✓ | | |
| Meldung von schwerwiegenden Vorfällen | ✓ | | |
| Aufbewahrung von erzeugten Protokollen | ✓ | | |
| Sofern relevant: Datenschutz-Folgenabschätzung | ✓ | | |
| Zusammenarbeit mit zuständigen nationalen Behörden | ✓ | | |
| Recht auf Erläuterung der Entscheidungsfindung im Einzelfall | ✓ | | |
| Informationspflichten gegenüber der Arbeitnehmerinnen-Vertretung<br>*sofern Arbeitgeberin Hochrisiko-KI-Systeme am Arbeitsplatz einsetzt* | ✓ | | |

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen (Kurz-Überblick)

**Gefahren bei der Nutzung / Bedrohungen durch KI**

Gegenmaßnahmen

Fazit

# Gefahren mittels KI
## (organisatorische Probleme)

Schatten-IT durch KI aktuell verstärkt

Rechtliche Probleme [auch beim Training der KI]

- Urheberrechte
- Datenschutz
- Geschäftsgeheimnisse (eigene / fremde)
- Sonstige: Markenrecht, UWG, AGG etc.

Wichtige Frage: Werden die eingegebenen Daten auch für das Training der KI genutzt?

# Gefahren mittels KI
## (Probleme bei der eigenen Nutzung)

Bias, Halluzinationen oder andere **Integritätsprobleme** durch

- veraltete oder „vergiftete“ Datenbasis
- schlechte Modell-Auswahl
- Schwache Daten/fehlende Qualität

Problematische **Nachvollziehbarkeit** der Ergebnisse

**Anker-Effekt** bei der Nutzung der Ergebnisse

# Gefahren bei Online-Services
## (KI auf der eigenen Website)

(Indirect) Prompt Injection

- Manipulation des Chatbots
- Manipulation durch versteckte Prompts

Poisening Attacks

Lieferketten-Sicherheitslücken

Rechtliche Probleme durch Non-Compliance

# Gefahren mittels KI

## (Probleme durch die Nutzung durch Dritte/Fremde)

- Unterstützung bei Schwachstellensuche
- Schnellere Analyse von Patches (Erhöhung von „Zero Days“?)
- Re-Identifizierung von anonymisierten Daten
- Verbesserung / Individualisierung von Malware
- Wissenssammlung und -aufbereitung bei Cyberangriffen
- Fake-News / Hoaxes
- Professionalisierung von Sc
- Ausnutzung hoher Rechte
- Schwachstellensuche und

Gefahren der Cybersecurity steigen durch KI

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen (Kurz-Überblick)

Gefahren bei der Nutzung / Bedrohungen durch KI

**Gegenmaßnahmen**

Fazit

# Gegenmaßnahmen



## KI-Strategie entwickeln

- Vermeidung von Schatten-KI
- KI-Kompetenz aufbauen
- Tipp: Beginn mit kleinen Projekten mit Quickwins

## Etablierung einer Fehler- und Sicherheitskultur

- Lessons Learned statt Bestrafung
- Sicherheit (inkl. Datenschutz) als Teil der Unternehmensstrategie
- Top-Down-Ansatz

# Gegenmaßnahmen



Sensibilisierung von

- Mitarbeitenden und (!!)
- Management

Anpassung der ISMS/DSMS

- Risikobewertung
- Organisation anpassen (z. B. BCM)
- aufbauende Prozesse (z. B. Codewords, 4-Augen-Prinzip)

# Gegenmaßnahmen

## IT-Sicherheitsmaßnahmen

- Patchmanagement
- Aufbau einer resilienten IT-Infrastruktur
- Verbesserung der Angriffserkennung
- Multi-Faktor- Authentifizierung
- eigene Nutzung der KI für Verteidigungsmaßnahmen (z. B. Erkennung von Bedrohungen und Schwachstellen).

# Gegenmaßnahmen



## Schutz der KI

- Sicherstellung der Qualität und der Integrität der Trainingsdaten
- Schutz der Modelle vor Diebstahl und Manipulation
- Durchführung von umfassenden Tests
- Sorgfältige Auswahl des KI-Systems und des betreibenden Unternehmens
- etc.

***beachte:***
*Da KI noch „neue“ Technologie ist, sind Schutzmaßnahmen z. T. noch schwierig umsetzbar (beispielsweise bei Injection)*

# Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen (Kurz-Überblick)

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

# Fazit

- KI betrifft alle Unternehmen, egal ob sie es aktiv nutzen
- „Wettrüsten“ zwischen den KI-Anwendern (Angreifer/Verteidiger)
- Risikobewertung und Anpassung des ISMS
- Etablierung einer KI-Strategie
- operativ: Überblick über (KI-) Systeme
  - Betrachtung (aller) rechtlicher Implikationen
  - Durchführung einer Risikobewertung

CHANGING PERSPECTIVE
BLICKWINKEL ZUKUNFT
Austrian IT & CONSULTANTS DAY 2025

# FRAGEN??

**thoffmann@uimc.at**

**Folgen Sie mir auf LinkedIn:**
**www.linkedIn.com/in/tim-hoffmann-uimc**