Präsentation
von René Kost

Ergänzend zur Projektarbeit
im Rahmen der
IHK-Zertifizierung:

IHK Cyber Security Advisor

Abgenommen durch

MIND_Refined
DIGITAL EDUCATION FOR A DIGITAL WORKING ENVIRONMENT

Zertifiziert durch

IHK Cottbus

IHK Cottbus

->End

Nutzung von KI – Systemen bei MDTP AG

René Kost

MDTP AG
Maschine Digital Twin Produktion

->End

Bildquelle: https://www.netzpiloten.de/die-besten-ki-tools-chatgpt-nightcafe-mehr/

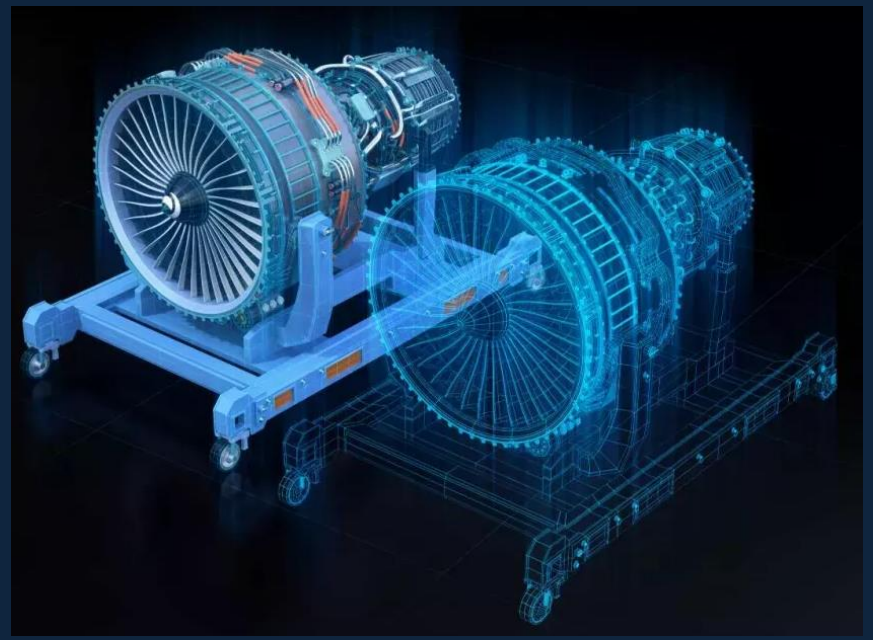
Agenda


Einführung
Risiken
Schutzbedarfs-
feststellung
Richtline
Fragen

->End

MDTP AG

Firmen Profil
- Kleines Unternehmen
- Hauptkunden sind im Maschinenbau und Produktion
- Wachstumspotentzial



Das Kerngeschäft ist die Prozessoptimierung mittels eines digitalen Zwillings.

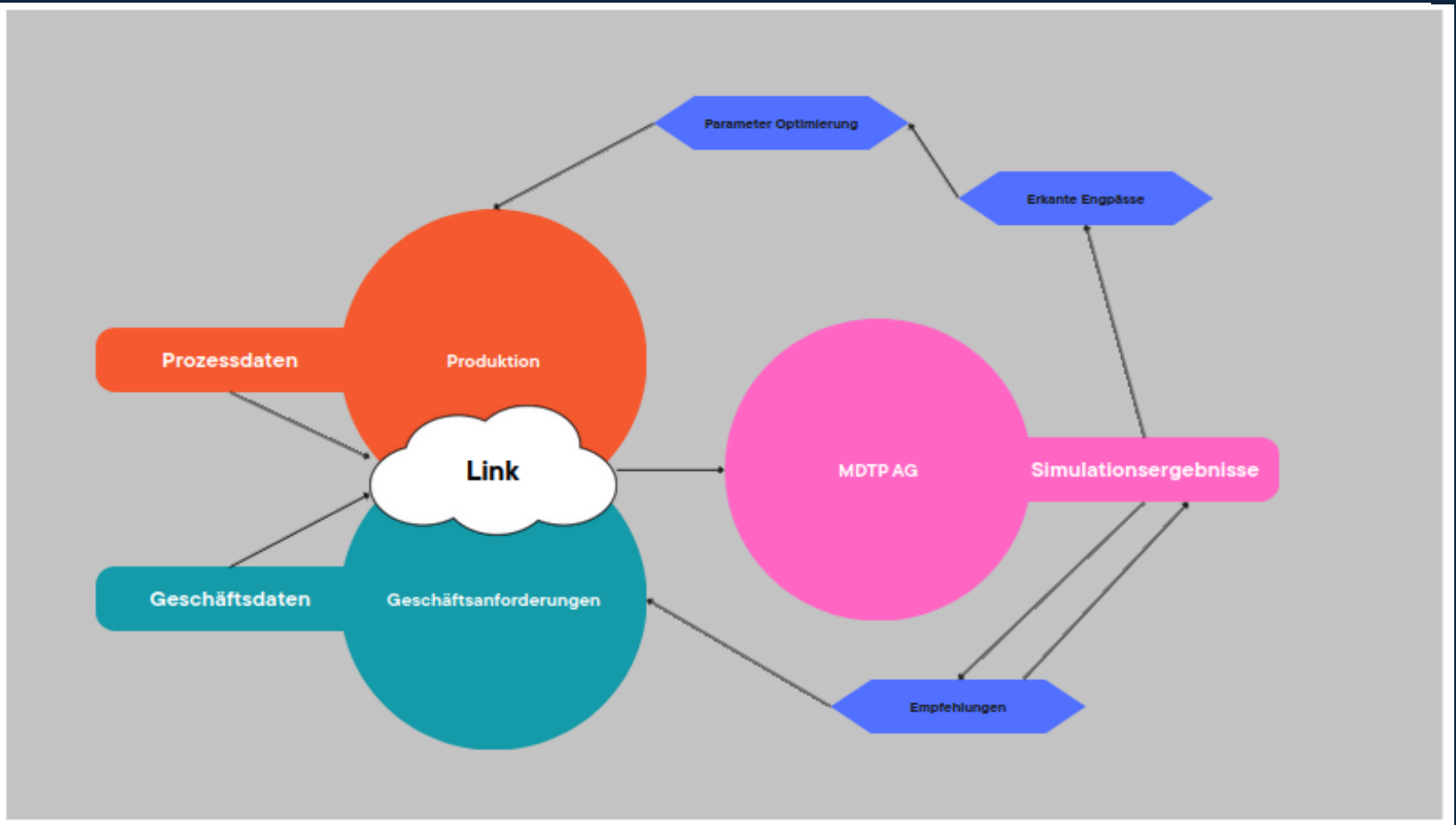
Was ist ein Digitaler Zwiling?

Nutzung von KI bei MDTP AG

->End

Bildquelle: https://www.automation-next.com/future-tech/damit-der-digital-twin-nicht-zum-boesen-zwilling-wird-768.html

Was ist ein Digitaler Zwiling?



Nutzung von KI bei MDTP AG

MDTP AG
Maschine Digital Twin Produktion

Richtline - Nutzung KI

Warum?

Leichtsinniger Umgang mit KI

Unternehmensinformationen bei ChatGPT entdeckt!!

Informationssicherheit: Schutz vor Datenabfluss und -lecks

Unterschiedliche Anforderunge an Abteilungen

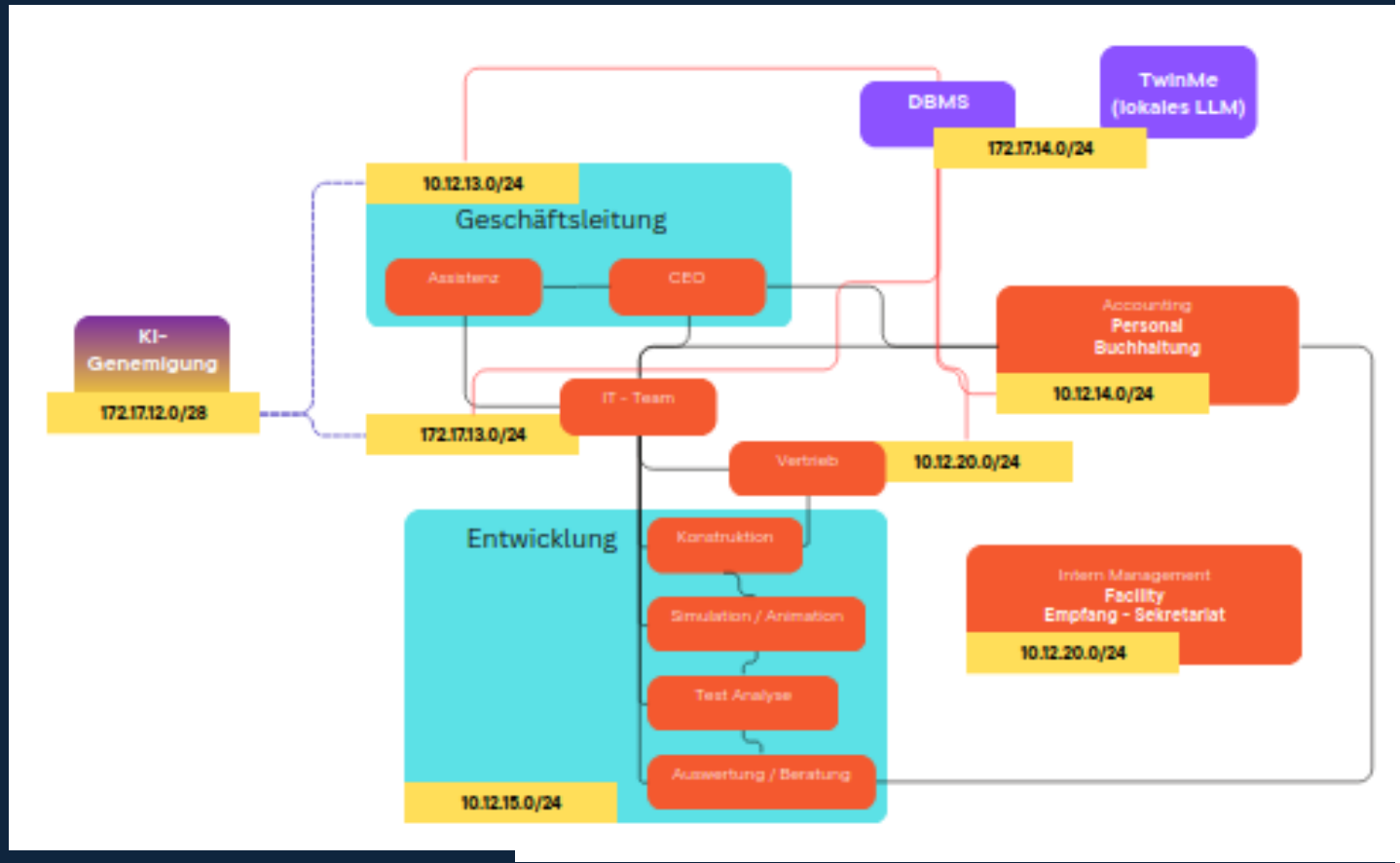
Anforderung an IT - Infrastrucktur

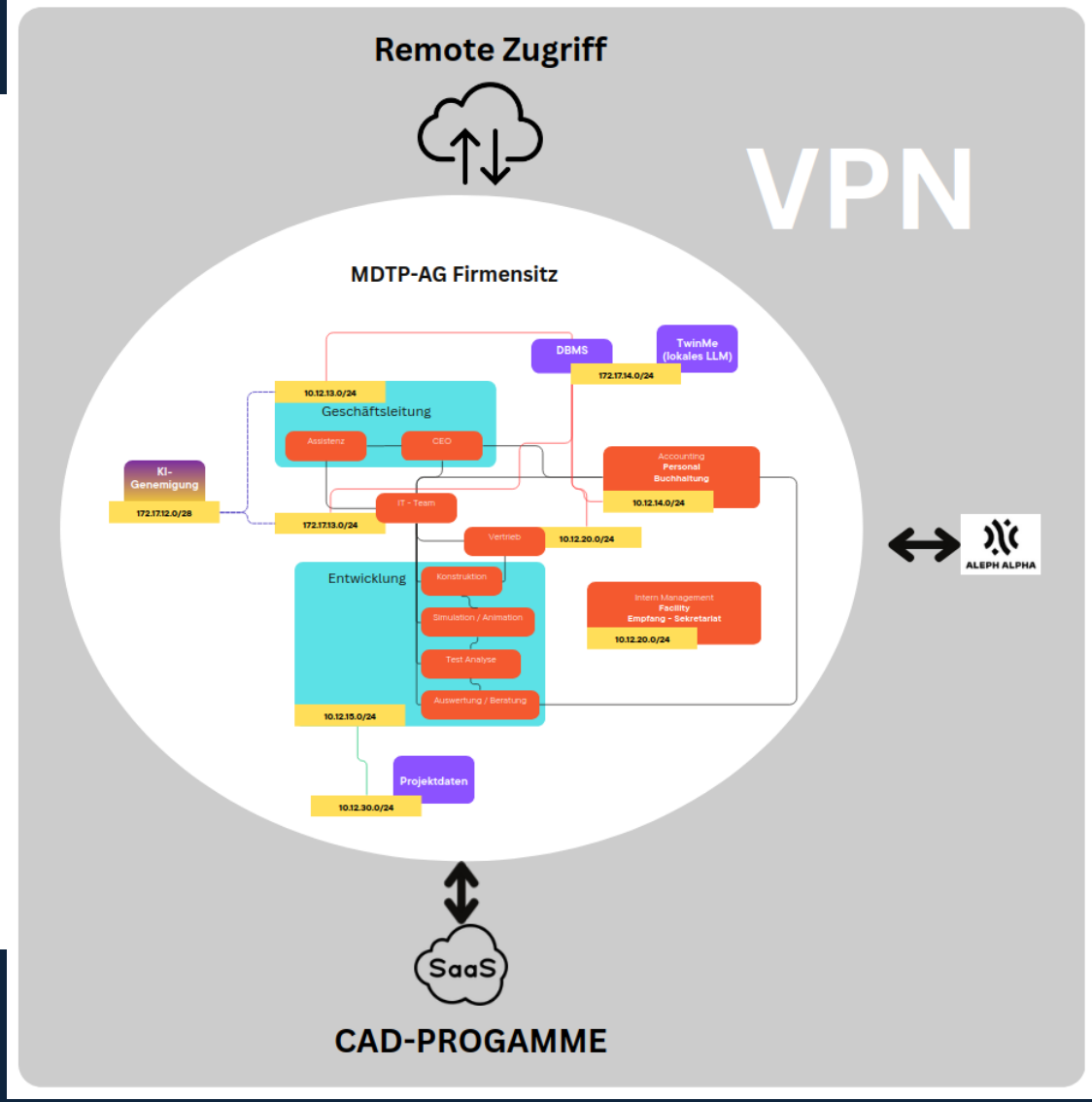
Schutzbedarfs-feststellung

->End

Bildquelle: https://t3n.de/news/ki-chatbots-als-psychologen-darum-ist-chatgpt-ein-schlechter-therapeut-1696037/

Oragnigramm MDTP AG

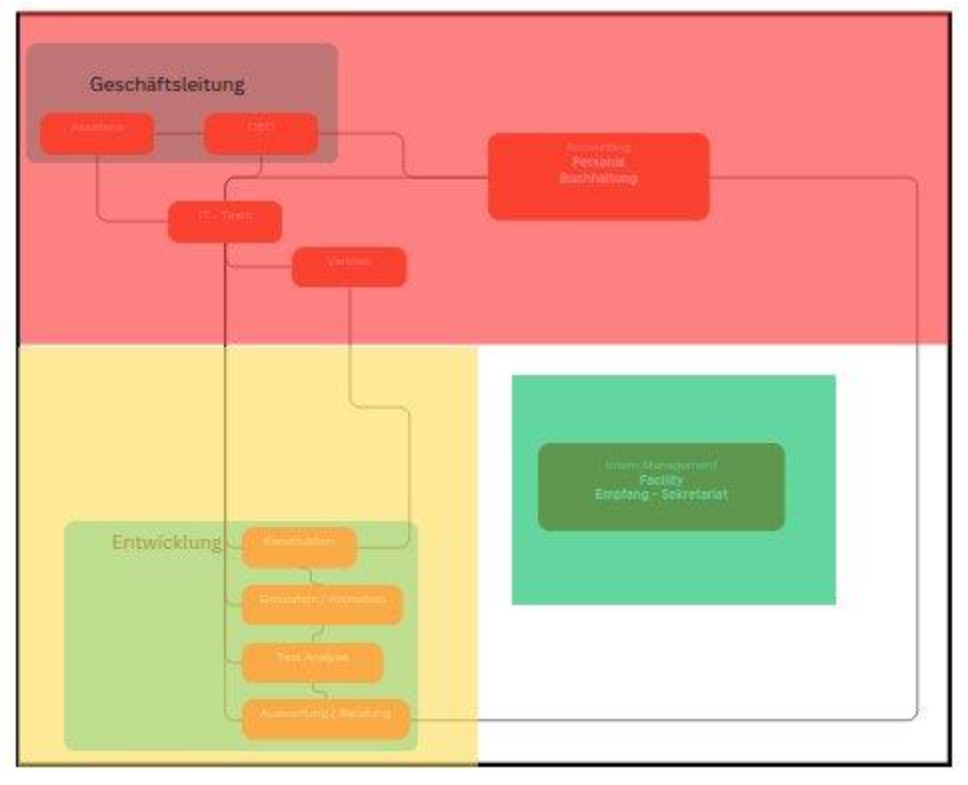




Nutzung von KI bei MDTP AG

->End

Schutzbedarfsfeststellung



Sehr hoher Schutzbedarf: Sensibelste Unternehmensdaten mit gesetzlichem Schutzanspruch

Hoher Schutzbedarf:
einzelne Projetktdaten

Normaler Schutzbedarf: Kaum oder kein Zugriff auf sensible Daten

⬇

Bestandsanalyse Form Soll / IST

⬇

Risiken

Bestandsanlyse Soll / Ist

Pflicht	Soll	Ist	
Schulungen	4	4	Nutzung // Verwendung
Datenkategorisierung + DLP	4	3	Pflege Trainingsdaten
Reglementierung der KI - Dienste	4	3	Ergänzung und Plfege (neu KI's)
Prozesshoheit	4	4	Plausibilitätschecks
Audit	4	4	Check erlaubter KI's

Risikoanalyse

Stufe	Bezeichnung	Auswirkung (prozentual)	Finanzieller Schaden (EUR)
1	Sehr niedrig	0 % – 1,5 %	bis 15.000 €
2	Niedrig	1,6 % – 5 %	bis 50.000 €
3	Akzeptabel	5,1 % – 9 %	bis 90.000 €
4	Mittel	9,1 % – 20 %	bis 200.000 €
5	Hoch	20,1 % – 60 %	bis 600.000 €
6	Sehr hoch	60,1 % – 100 %	über 600.000 €

Schwachstelle	SLE	ARO	ALE
Datenablfluss	3	24	72
Seriösität der Daten	4	12	48
Fehlende Schulungen	1	600	600

Bsp.

900.000Euro

Regelungen der Richtline

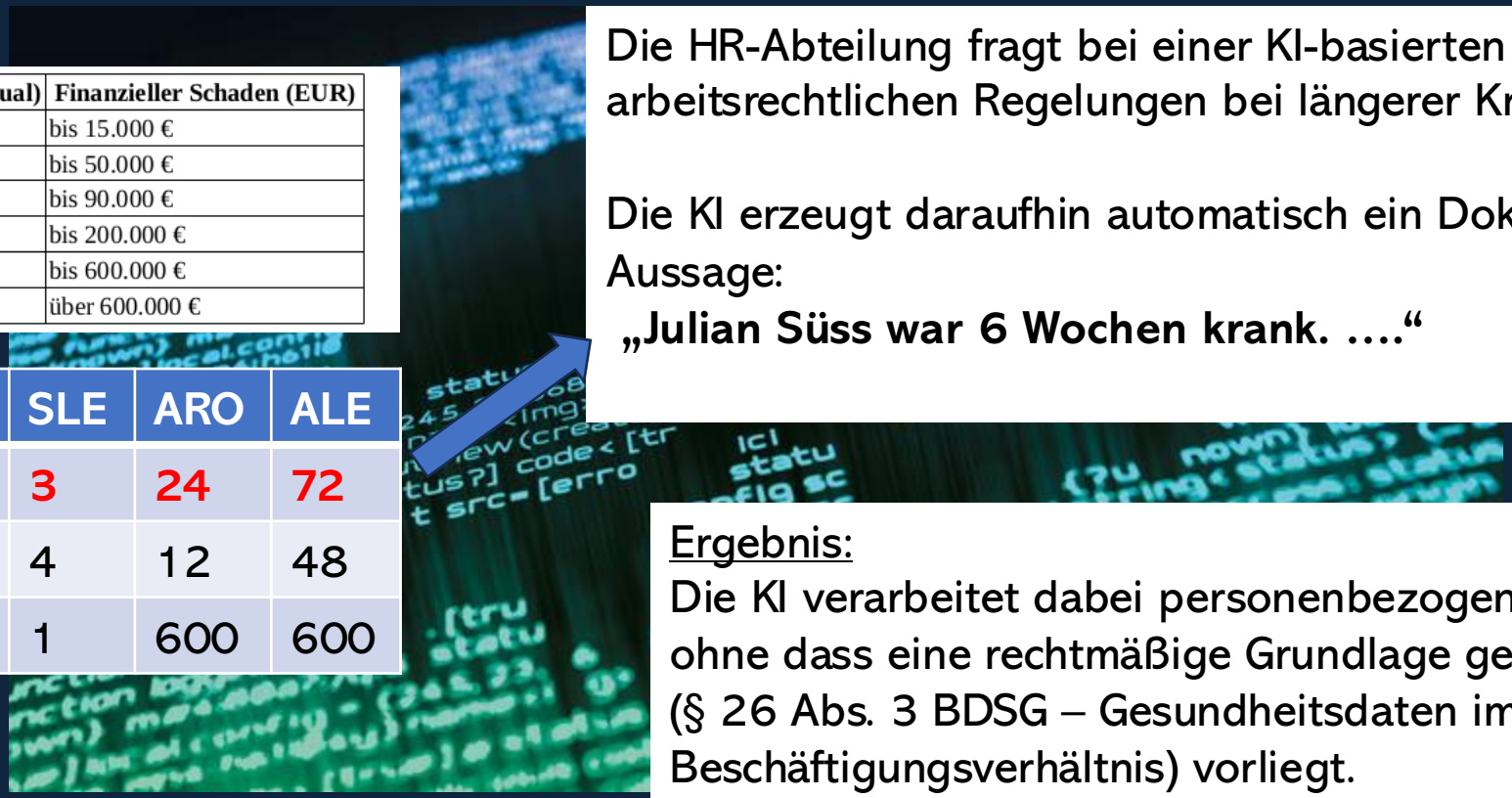
->End

Bildquelle: https://research.hisolutions.com/2025/03/ki-agenten-anfaellig-fuer-einfache-gefaehrliche-angriffe/

Bsp. Risikoanalyse

Stufe	Bezeichnung	Auswirkung (prozentual)	Finanzieller Schaden (EUR)
1	Sehr niedrig	0 % – 1,5 %	bis 15.000 €
2	Niedrig	1,6 % – 5 %	bis 50.000 €
3	Akzeptabel	5,1 % – 9 %	bis 90.000 €
4	Mittel	9,1 % – 20 %	bis 200.000 €
5	Hoch	20,1 % – 60 %	bis 600.000 €
6	Sehr hoch	60,1 % – 100 %	über 600.000 €

Schwachstelle	SLE	ARO	ALE
Datenablfluss	3	24	72
Seriösität der Daten	4	12	48
Fehlende Schulungen	1	600	600



Die HR-Abteilung fragt bei einer KI-basierten Anwendung nach arbeitsrechtlichen Regelungen bei längerer Krankheit.

Die KI erzeugt daraufhin automatisch ein Dokument mit der Aussage:
„Julian Süss war 6 Wochen krank. ….“

Ergebnis:
Die KI verarbeitet dabei personenbezogene Gesundheitsdaten, ohne dass eine rechtmäßige Grundlage gemäß DSGVO und BDSG (§ 26 Abs. 3 BDSG – Gesundheitsdaten im Beschäftigungsverhältnis) vorliegt.

➡ Verstoß gegen die Datenschutz-Grundverordnung (DSGVO) und das Bundesdatenschutzgesetz (BDSG).

Weiter

ARO = 2 mal im Monat

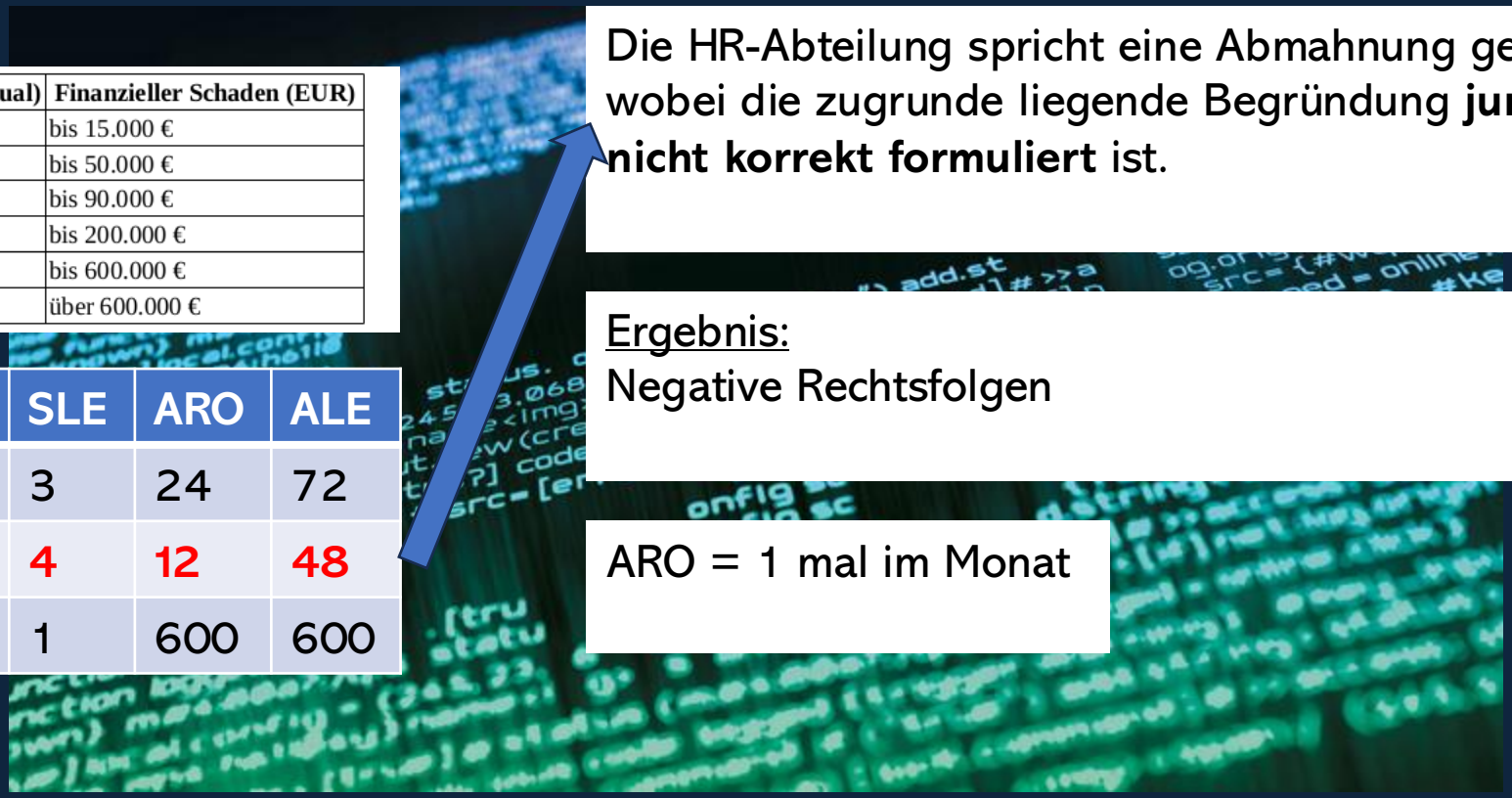
->End

Bildquelle: https://research.hisolutions.com/2025/03/ki-agenten-anfaellig-fuer-einfache-gefaehrliche-angriffe/

Bsp. Risikoanalyse

Stufe	Bezeichnung	Auswirkung (prozentual)	Finanzieller Schaden (EUR)
1	Sehr niedrig	0 % – 1,5 %	bis 15.000 €
2	Niedrig	1,6 % – 5 %	bis 50.000 €
3	Akzeptabel	5,1 % – 9 %	bis 90.000 €
4	Mittel	9,1 % – 20 %	bis 200.000 €
5	Hoch	20,1 % – 60 %	bis 600.000 €
6	Sehr hoch	60,1 % – 100 %	über 600.000 €

Schwachstelle	SLE	ARO	ALE
Datenablfluss	3	24	72
Seriösität der Daten	**4**	**12**	**48**
Fehlende Schulungen	1	600	600



Die HR-Abteilung spricht eine Abmahnung gegen **Julian Süss** aus, wobei die zugrunde liegende Begründung **juristisch und sachlich nicht korrekt formuliert** ist.

Ergebnis:
Negative Rechtsfolgen

ARO = 1 mal im Monat

Weiter

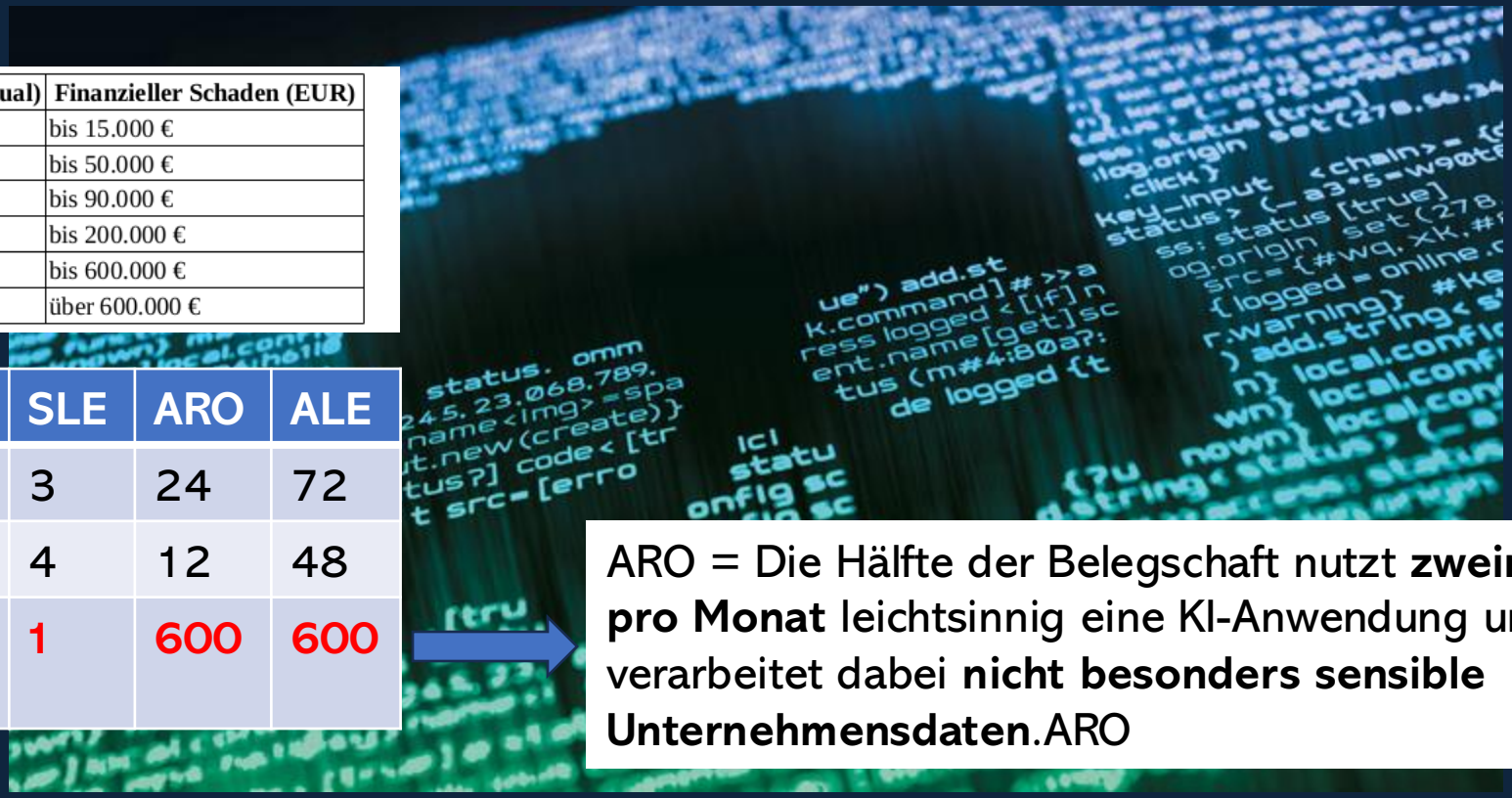
->End

Bildquelle: https://research.hisolutions.com/2025/03/ki-agenten-anfaellig-fuer-einfache-gefaehrliche-angriffe/

Bsp. Risikoanalyse

Stufe	Bezeichnung	Auswirkung (prozentual)	Finanzieller Schaden (EUR)
1	Sehr niedrig	0 % – 1,5 %	bis 15.000 €
2	Niedrig	1,6 % – 5 %	bis 50.000 €
3	Akzeptabel	5,1 % – 9 %	bis 90.000 €
4	Mittel	9,1 % – 20 %	bis 200.000 €
5	Hoch	20,1 % – 60 %	bis 600.000 €
6	Sehr hoch	60,1 % – 100 %	über 600.000 €

Schwachstelle	SLE	ARO	ALE
Datenablfluss	3	24	72
Seriösität der Daten	4	12	48
Fehlende Schulungen	**1**	**600**	**600**

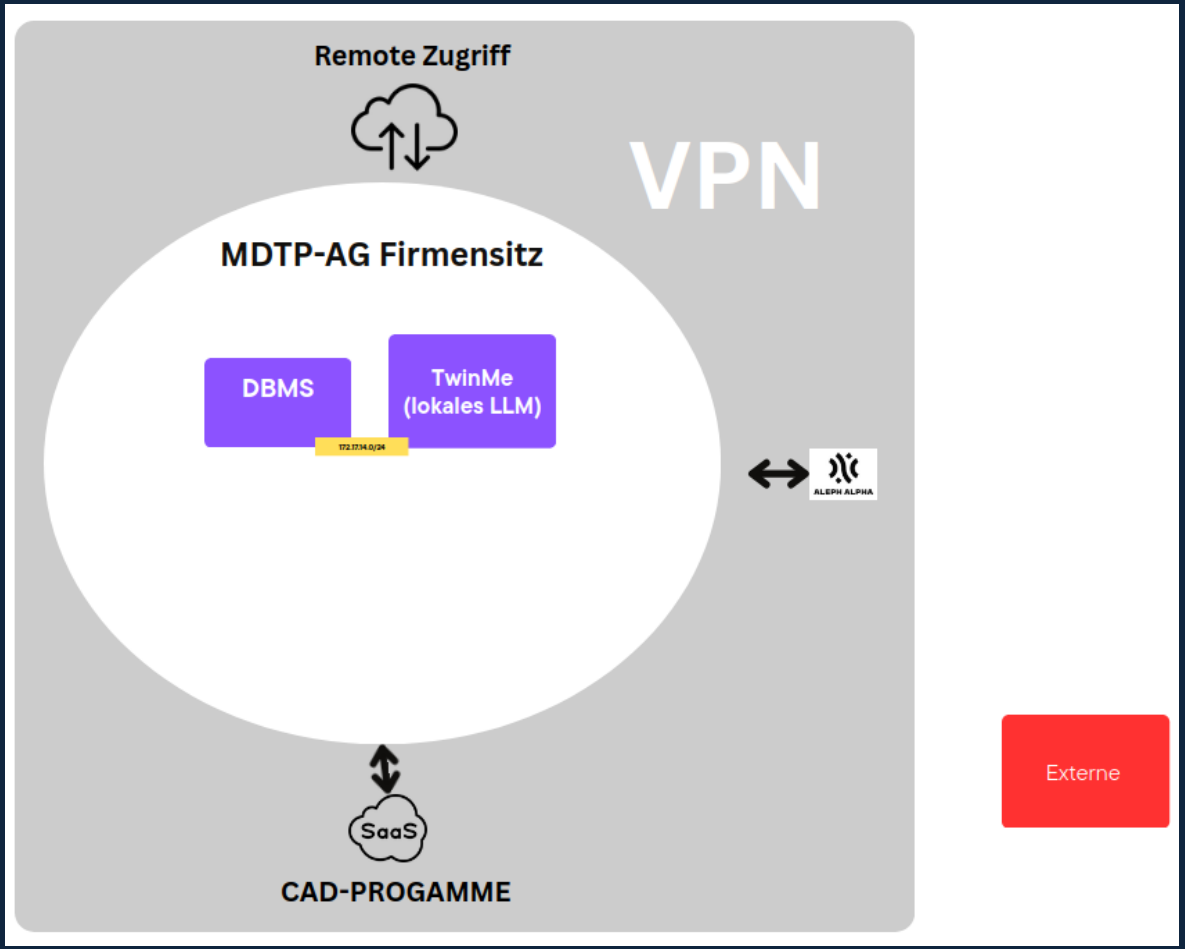


ARO = Die Hälfte der Belegschaft nutzt **zweimal pro Monat** leichtsinnig eine KI-Anwendung und verarbeitet dabei **nicht besonders sensible Unternehmensdaten**.ARO

Bildquelle: https://research.hisolutions.com/2025/03/ki-agenten-anfaellig-fuer-einfache-gefaehrliche-angriffe/

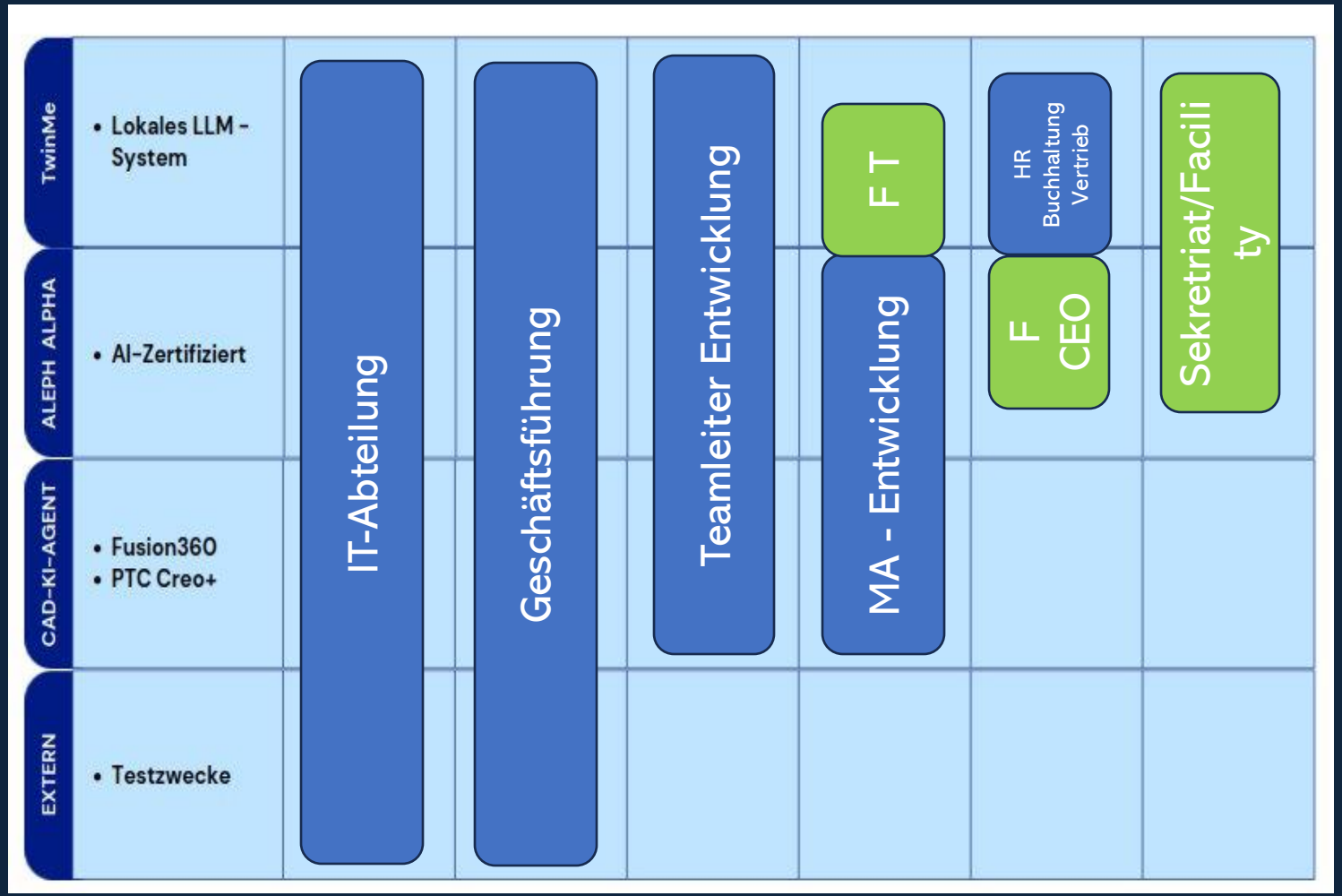
Grundlegenden Nutzung

Weche LLM sind erlaubt ?



Grundlegenden Nutzung

Nutzungsrechte ?



->End

Weitere Regelungen

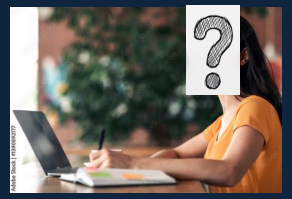
Begrenzte Nutzungszeiten von 8:00 – 20:00 Uhr

Dokumentationspflicht

Anonymisierung min. Verschleierung

juristischen Bewertung

Grundlegenden Regelungen



Dokumentationspflicht und Nutzungsvorausetzung

Notation im Entwicklungslog

Nur bis Layer 2

Verschleierung um den Grössenfaktor 11

Die Dokumentation ist unter:
/BUSINESS/AI_RULE/USE/<Name_Mitarbeiter>.md
im Markdown-Format zu erstellen.

Bearbeitetes Dokument: <Name>
Datum: <Datum>
Verwendetes KI-System: <KI-System>

Bilder mit *KI-generiert* zu kennzeichnen

Nutzung von *PromptSecurity*

Promting

Charles Reade
"Achte auf deine Gedanken, …. , denn er wird dein Schicksal."

Deine Promts werden dein Unternehmesschicksal
René Kost

Prompting

Verboten:
- Zugangsdaten
- Konfigurarationen
- Personen bezogene und Kundendaten
- Verhaltenskodex

Erlaubt:
- Anonymisierte Inhalte
- Öffentlich zugängliche Informationen

Fehleranalyse
->End

Fehleranalyse

Jedes Verwendete Ergebniss muss geprüft werden!

Dokumentation unter:
/BUSINESS/AI_RULE/ERROR/<DATUM_ABTEILUNG_FEHLERNAME>.md

Prompt: *<Prompt>*
KI-Antwort: *<Ki-Ergebnis>*
Fehlerbeschreibung: *<Fehlerbeschreibung>*

TwinMe
- Validierung der Trainingsdaten
- KeywordError

20% -> Datenschutzbeaugtagte kann das Modul sperren

Verantwortlichkeiten

- IT-Admins
  - Kontinuierliche Weiterentwicklung der lokalen LLM
  - Einpflegen der Trainingsdaten
  - Bug-Bounty-Programm
  - Installation und Updates der Filtersysteme

- Datenschutzbeauftragte
  - Updates zur Sicherheitslage der verwendeten KI-Systeme
  - Genehmigung von Trainingsdaten und neuen KI-Systemen
  - Entzug von Rechten

- **Personalabteilung**
  - **Einhaltung von Schulungen**

Dokumentenverantwortung (CEO/ISM)

Ansprechpartner bei Problemen und Rückfragen; Dokumentenprüfung

Schulungen

Für die NUTZUNG der genemigten KI-Systeme:
MDTP-SW-007-021 [KI-Grundlagen]

Für die VERWENDUNG genemigten KI-Systeme:
MDTP-SW-011-[KI-Bias-Erkennung und ethischen Aspekten]

Zeit für Fragen

Digitaler Zwiling
IT-Strucktur
Dokumentation
Fehleranalyse
Schulungen

MDTP AG
Warum KI-Richtline
Schutzbedarf
Risiken
Regelungen
Prompting
Verantwortlichkeiten

KI-generiertes Bild:ChatGpt