

# KI-Angriffserkennung:

## Mit künstlicher Intelligenz und Machine-Learning Angriffe aufspüren



Prof. Dr. Kai-Oliver Detken  
DECOIT GmbH & Co. KG  
Fahrenheitstraße 9, D-28359 Bremen  
<https://www.decoit.de>  
[detken@decoit.de](mailto:detken@decoit.de)

- **IT-Consulting:** ganzheitliche sowie herstellerneutrale Beratung
- **System Management:** Optimierung technischer Arbeitsabläufe, Integration von Hersteller- oder Open-Source-Lösungen in vorhandene Umgebungen
- **Software-Entwicklung:** Entwicklung von Individualsoftware, Anpassung bestehender Open-Source-Software an Kundenbedürfnisse
- **IT-Forschungsprojekte:** innovative IT-Lösungen
- **Produktentwicklung:** innovative Produkte auf Basis von F&E-Projekten

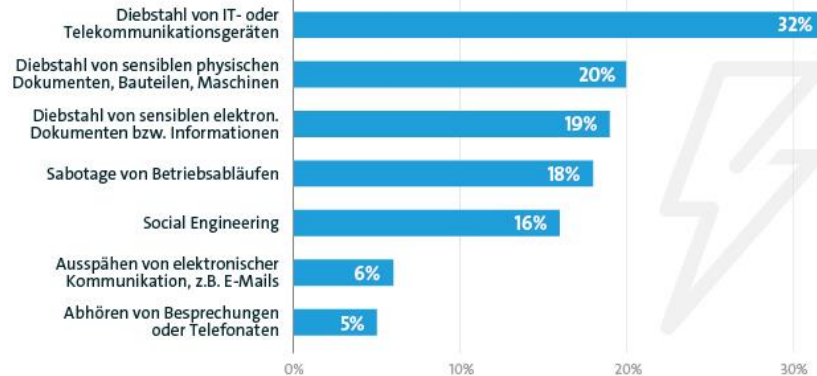


## Datenklau, Spionage, Sabotage: Zwei Drittel der Industrie betroffen



Basis: Alle befragten Industrieunternehmen (n=504)  
Quelle: Bitkom Research

### Die häufigsten Delikte



**22,35 Mrd. Euro Schaden pro Jahr**

bitkom

Quelle: <https://ap-verlag.de/industrie-im-visier-von-cyberkriminellen-und-nachrichtendiensten/20754/>

## Ransomware

ist weiterhin die größte Bedrohung.

**2** Ransomware-Angriffe auf Kommunalverwaltungen oder kommunale Betriebe wurden durchschnittlich pro Monat bekannt.

**68** erfolgreiche Ransomware-Angriffe auf Unternehmen wurden bekannt.

**15** davon richteten sich gegen IT-Dienstleister.




**2.000** Mehr als Schwachstellen in Softwareprodukten (15 % davon kritisch) wurden im Berichtszeitraum durchschnittlich im Monat bekannt. Das ist ein Zuwachs von 24 %.

**Eine Viertelmillion** neue Schadprogramm-Varianten wurden durchschnittlich an jedem Tag im Berichtszeitraum gefunden.

**66%** aller Spam-Mails im Berichtszeitraum waren Cyberangriffe: 34 % Erpressungsmails, 32 % Betrugsmails

**84%** aller betrügerischen E-Mails waren Phishing-E-Mails zur Erhebung von Authentisierungsdaten, meist bei Banken und Sparkassen.

Top-3-Bedrohungen je Zielgruppe:

| Gesellschaft  | Wirtschaft   | Staat und Verwaltung   |
|---|--|--|
| <br><b>Identitätsdiebstahl</b><br>Sturmtrommel<br>Phishing | <br><b>Ransomware</b><br>Abhängigkeit innerhalb der IT-Supply-Chain<br>Schwachstellen, offene oder falsch konfigurierte Online-server | <br><b>Ransomware</b><br>APT<br>Schwachstellen, offene oder falsch konfigurierte Online-server |

Rund **21.000** infizierte Systeme wurden täglich im Berichtszeitraum erkannt und vom BSI an die deutschen Provider gemeldet.

Durchschnittlich rund **775** E-Mails mit Schadprogrammen wurden an jedem Tag im Berichtszeitraum in deutschen Regierungszentralen abgefangen.

**370** Webseiten wurden im Durchschnitt an jedem Tag des Berichtszeitraums für den Zugriff aus den Regierungszentralen gesperrt. Der Grund: Die Seiten enthielten Schadprogramme.

**6.220** 2022  
**5.100** 2021

**7.120** Teilnehmer hatte die Allianz für Cybersicherheit im Jahr 2023.

Deutschland Digital-Sicher-BSI

- Die Bedrohungslage im Bereich der Cyber-sicherheit ist weiterhin von einer hohen Dynamik geprägt
- Die rasante Entwicklung im Bereich der Künstlichen Intelligenz zeigt, wie schnell technische Neuerungen fortschreiten können
- Nach wie vor bleiben Angriffe mit Ransomware die größte Bedrohung für die Cybersicherheit

Quelle: <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2023.html>

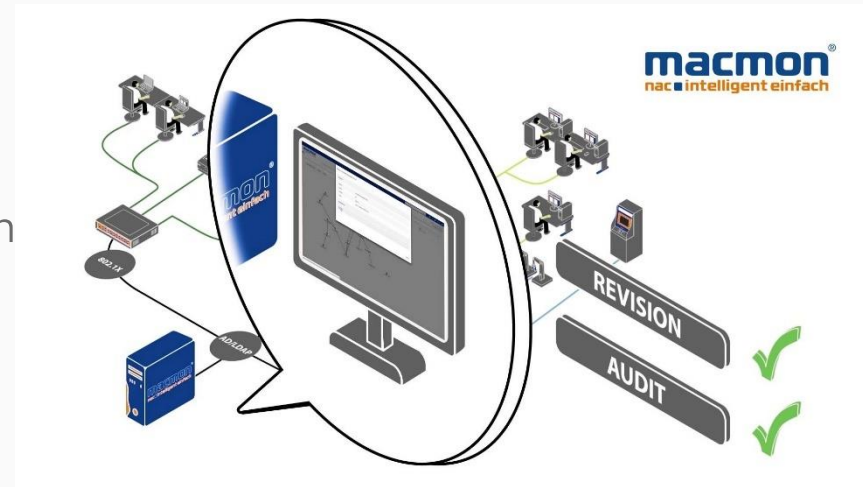
- Evolution der Überwachungs- und Regulierungssysteme:
  - **Netzmonitoring**: Überwachung der Verfügbarkeit und Netzdokumentation
  - **Network Access Control (NAC)**: Überwachung der Zugangskontrolle und Endgeräte-Dokumentation
  - **Security Information and Event Management (SIEM)**: Überwachung der IT-Sicherheit und Korrelation der Ereignisse (Vorfälle)
  - **Endpoint Detection and Response (EDR)**: Überwachung von Endgeräten bzgl. IT-Sicherheit und Anomalien (AV-Weiterentwicklung)

- Ziel: Überwachung von Services und Serversystemen sowie Sammeln von Verfügbarkeitsstatistiken
- Aufgaben:
  - Einbindung von Netzwerk- und Serverkomponenten
  - Überwachung von Services (Diensten)
  - Eskalationsmanagement bei Alarmmeldungen (SMS, E-Mail)
  - Zusammenfassung von Alarmmeldungen
  - Unterscheidung unterschiedlicher Prioritäten

The screenshot displays the DECOIT network monitoring interface. On the left is a sidebar with navigation options like 'Tactical Overview', 'Quicksearch', 'Dashboards', 'Views', and 'WATO - Quickaccess'. The main area is titled 'Main Overview' and contains several panels:

- HOST STATISTICS:** Shows Up (762), Down (7), Unreachable (0), and In DownTime (0).
- SERVICE STATISTICS:** Shows OK (33082), In DownTime (0), On Down host (0), Warning (37), Unknown (0), and Critical (17).
- HOST PROBLEMS (UNHANDLED):** A table listing issues like 'carsv0142ldap' and 'mucap0213san' with their respective ages and status details.
- SERVICE PROBLEMS (UNHANDLED):** A table listing service issues such as 'mucsv0456sql' (DB2 Backup), 'carsv0234ldap' (System Time), 'mlsv0532sql' (MySQL Blocked Sessions), 'mucsv0443sql' (DB2 Tablespace), 'lvsrv0413jvm' (JVM CBF Threads), 'mucsv1228lic' (Citrix Terminal Licensing), and 'lvsrv0413jvm' (JVM LOGSERVER Threads).
- EVENTS OF RECENT 4 HOURS:** A log of recent events with columns for Time, Alias, Service, and Output, showing various warnings and errors.

- Ziel: Zugangskontrolle von Systemen und Benutzern in Netzwerke
- Aufgaben:
  - Fremde Systeme erkennen
  - Auf Richtlinienkonformität überprüfen
    - Scan der installierten Programme
    - Scan der Sicherheitsupdates
  - Zugangsberechtigung erteilen oder verweigern
  - Verschieben von Systemen in bestimmte Netzwerke aufgrund der Richtlinien



- Ziel: Gesamtübersicht über den Sicherheitsstatus des Netzwerkes bieten
- Aufgaben:
  - Sammeln sicherheitsrelevante Informationen im Netzwerk
  - Bewerten dieser Informationen
  - Priorisierung der bewerteten Informationen
  - Meldungen über kritische Sicherheitslage geben
  - Handlungsempfehlungen bereitstellen

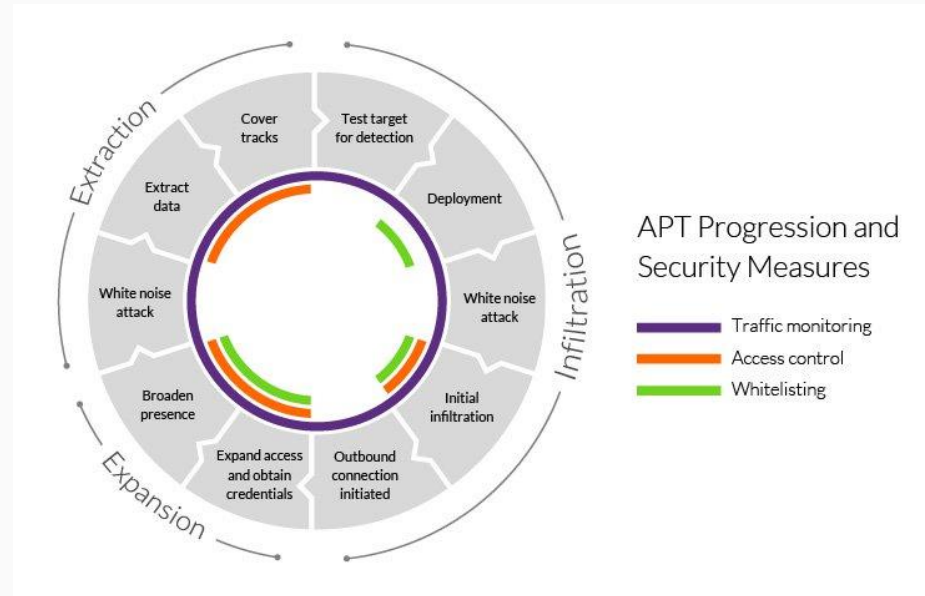




- Ziel: Überwachung von Endgeräten im Hinblick auf verdächtige Aktivitäten
- Aufgaben:
  - Konsolenbenachrichtigungen und -berichte
  - Leistungsfähige Analyse- und Reaktionsfunktionen
  - Sicherheitslücken auf Endgeräten erkennen und melden
  - Managed Services: Weitergabe an ein Security Operation Center (SOC)
  - Schutz gegen Advanced Persistent Threats (APT)

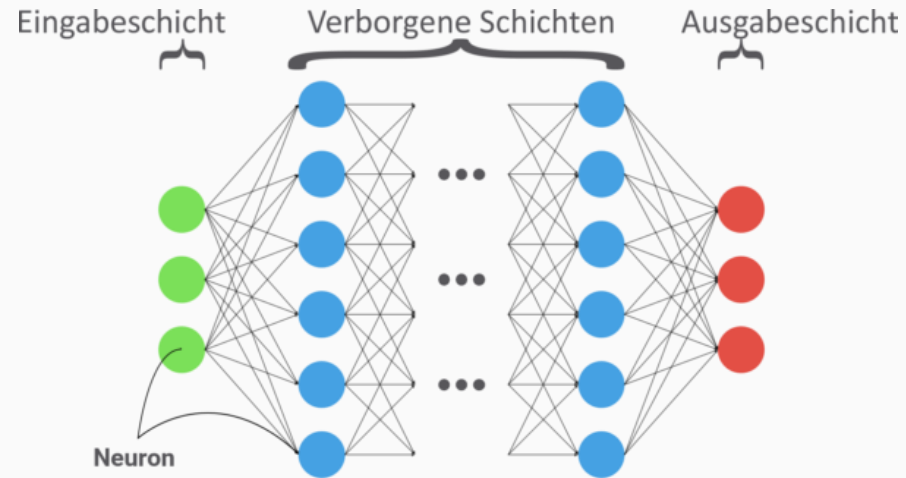


- APT ist ein komplexer, zielgerichteter Cyber-Angriff
- Der Angriff ist darauf ausgelegt über einen längeren Zeitraum hinweg unbemerkt im Zielsystem zu bleiben
- Schwachstellen des Zielsystems sollen auskundschaftet werden
- APTs sind schwer zu erkennen, da sie keinen unmittelbaren Alarm in Sicherheitssystemen auslösen
- Sie lassen sie sich durch die Nutzung von KI einfacher enttarnen, als durch den Sicherheitsexperten



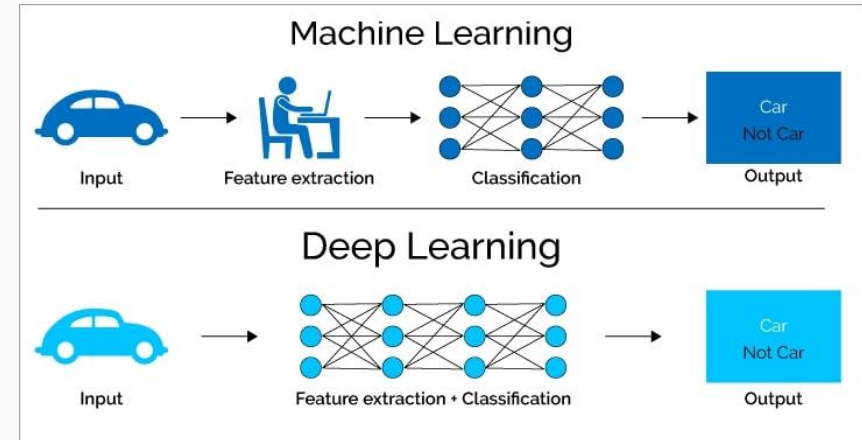
Quelle: <https://www.imperva.com/learn/application-security/apt-advanced-persistent-threat/>

- Generell spricht man bei Machine Learning (ML) von Verfahren, die es erlauben anhand bestimmter Merkmale Dinge zu klassifizieren und zu erwartende Ergebnisse zu extra- bzw. interpolieren
- Dafür werden Neuronale Netze (NN) verwendet, die es erlauben aus dem Dateninput über Synapsen-artige Verschaltungen wahrscheinliche Ergebnisse vorherzusagen
- Dabei gilt: je kleiner der Datensatz ist, umso unwahrscheinlicher ist die NN-Ausgabe korrekt!
- Das NN besteht aus einem Algorithmus, dem ein Entscheidungsbaum antrainiert werden kann (z.B. TensorFlow)
- Sie bilden die Grundlage für das Deep Learning



Quelle: <https://nativdigital.com/neuronale-netze/>

- Deep Learning ist eine spezielle Art des maschinellen Lernens, um Muster und Zusammenhänge in großen Datenmengen zu erkennen
- Deep Learning wird zunehmend in der Cybersicherheit eingesetzt, um fortschrittliche Angriffe wie APT-Angriffe zu erkennen, zu analysieren und abzuwehren:
  - Erkennen von Anomalien
  - Malware-Erkennung
  - Verhaltensanalyse
  - Vorausschauende Bedrohungsinformationen
  - Automatisierung von Sicherheitsmaßnahmen



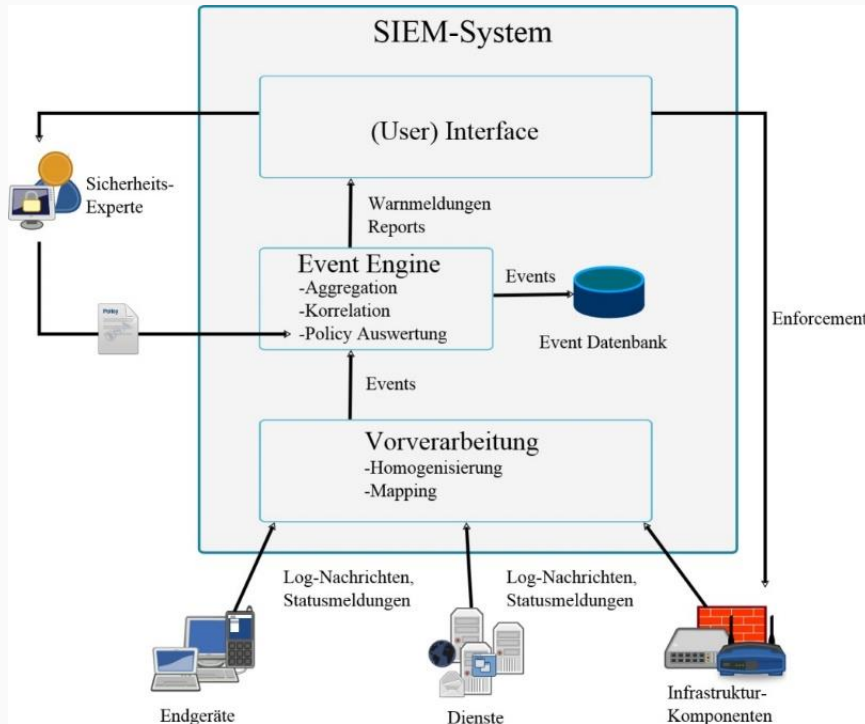
Quelle: <https://levity.ai/blog/difference-machine-learning-deep-learning>

- TensorFlow ist eine Open-Source-Plattform für maschinelles Lernen und Deep Learning, die von Google entwickelt wurde
- Sie bietet eine umfassende Sammlung von Tools, Bibliotheken und Ressourcen, die es Entwicklern ermöglichen, maschinelle Lernmodelle zu erstellen, zu trainieren und bereitzustellen
- TensorFlow wurde für eine Vielzahl von Anwendungen entwickelt, darunter Bild- und Spracherkennung, Natural Language Processing (NLP), Zeitreihenanalyse etc.
- SIEM-Systeme basieren hauptsächlich auf TensorFlow



# TensorFlow

- SIEM-Systeme nutzen KI-Algorithmen auf verschiedene Weise, um Angriffe zu erkennen:
  - **Verhaltensanalyse:** SIEM-Systeme können KI-Algorithmen verwenden, um das normale Verhalten von Benutzern, Geräten und Anwendungen in einem Netzwerk zu modellieren.
  - **Bedrohungsintelligenz:** KI-Algorithmen können dabei helfen, große Mengen von Bedrohungsdaten zu analysieren, um relevante Informationen zu identifizieren.
  - **Automatisierung von Reaktionen:** Moderne SIEM-Systeme integrieren oft automatisierte Reaktionen auf Bedrohungen.
  - **Erkennung von unbekanntem Bedrohungen:** KI-Algorithmen sind auch in der Lage, Anomalien zu erkennen, die auf bislang unbekannte Bedrohungen hinweisen könnten.
- Grundsätzlich werden nicht verschiedene Algorithmen genutzt, sondern verschiedene Datensätze verwendet.



- Log- und Netzwerkdaten sammeln (Big Data)
- Homogenisieren der Daten zu einer Datenbasis
- Aggregation verschiedener Events
- Trainieren des Normalverhaltens, um Anomalien mittels KI zu erkennen
- Bewerten der Vorfälle anhand eines Sicherheitsexperten und mittels KI

- Durch die CTI-Nutzung können potenzielle Gefahren erkannt und abgewendet werden
  - CTI beinhaltet die kontinuierliche Sammlung von Informationen über potenzielle Cyberbedrohungen aus verschiedenen Quellen
  - Die gesammelten Daten werden analysiert und bewertet
  - Schwachstellen und Angriffsmuster sollen besser erkannt werden
  - CTI-Plattformen erstellen Bedrohungsmeldungen, die von SIEM-Systemen eingelesen werden können (sog. Threat-Feeds)
- Die CTI-Datenbank dient dazu, bereits bekannte Bedrohungsindikatoren von außerhalb zu erhalten, um neue interne Angriffe besser identifizieren zu können
- Die gewonnenen Erkenntnisse werden in verständlicher Form zusammengefasst → Handlungsempfehlungen



- Auch leistungsfähige SIEM-Systeme sind nur so gut wie ihre Datenbasis
  - Ab wann ist ein Normalverhalten erreicht?
  - Wurden zum Anlernen des KI-Algorithmus echte Daten verwendet?
  - Schickt der Hersteller regelmäßig Threat-Feeds?
  - Ist ein Security Operation Center (SOC) vorgesehen?
- KI-Verfahren können helfen die Datenmenge zu beherrschen und die relevanten Informationen herauszufiltern
- Automatisierte Reaktionen sind allerdings in manchen Umgebungen mit Vorsicht einzusetzen

# Vielen Dank für die Aufmerksamkeit!



DECOIT GmbH & Co. KG  
Fahrenheitstraße 9  
D-28359 Bremen

<https://www.decoit.de>  
[info@decoit.de](mailto:info@decoit.de)

