German American Chambers of Commerce
Deutsch-Amerikanische Handelskammern

TeleTrusT Pioneers in IT security.
IT Security Association Germany

German-American IT Security Forum

IT Security Association Germany (TeleTrusT) in cooperation with GACC – German American Chamber of Commerce

San Francisco, 28.04.2025

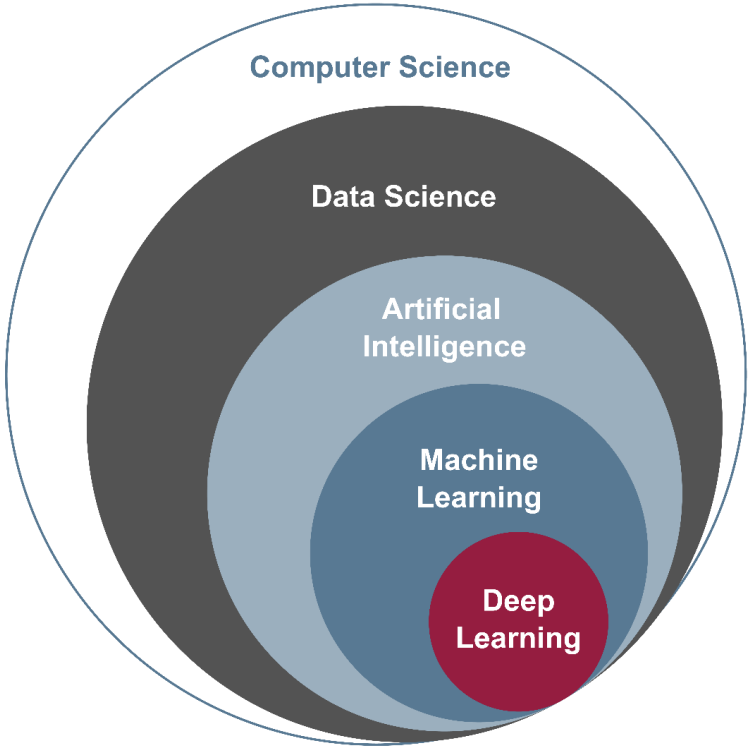
AI and IT security
→ more security, more threats

Prof. Dr. Norbert Pohlmann

Professor for Cyber Security
Director of the Institute for Internet Security – if(is)
Chairman of the board of the IT Security Association TeleTrusT
Member of the board of the Internet industry association eco.

Classification

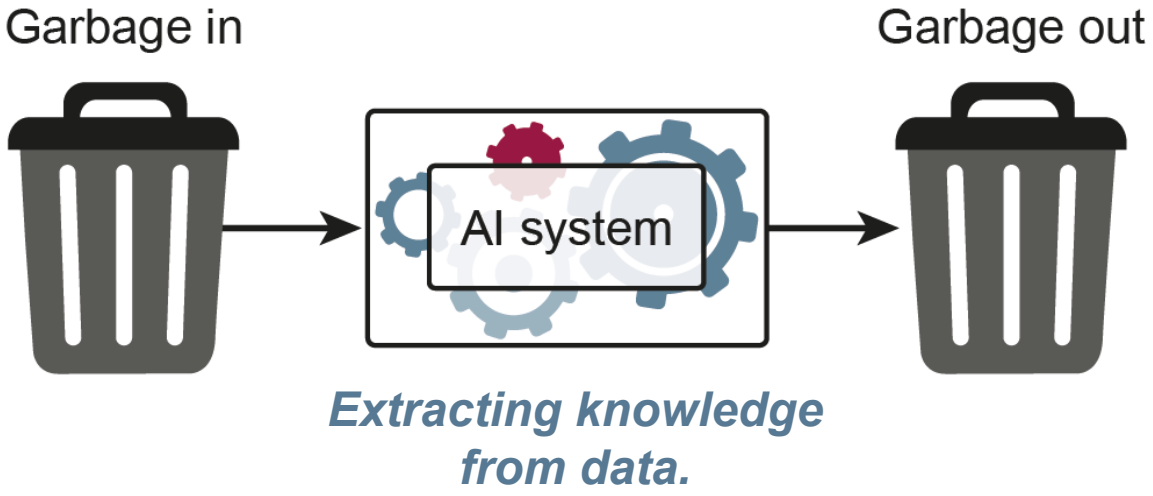
→ (Artificial intelligence) **Machine learning**



- **Data Science** generally refers to the **extraction of knowledge from data**.
- **Strong "Artificial Intelligence"** *(future)*
  is intended to automatically emulate **„human-like intelligence“**.
  Singularity, Artificial General Intelligence (AGI) - („machines“ improve themselves, are *„more intelligent“ than humans*)
- **Weak „artificial intelligence“** *(very successful today)*
  Machine learning (ML) is a term for the generation of knowledge from experience (in data) by computer.
- **Deep Learning** → Method of ML → **Improved results**
- **Large Language Modell (LLM)** Basic model for generative AI ***(stochastic parrot)***
- **Generative AI (GenAI)** *is a form of artificial intelligence that can produce text, images, code and various other content based on its training data*.

A principle of artificial intelligence
→ Garbage in – Garbage out

Paradigma



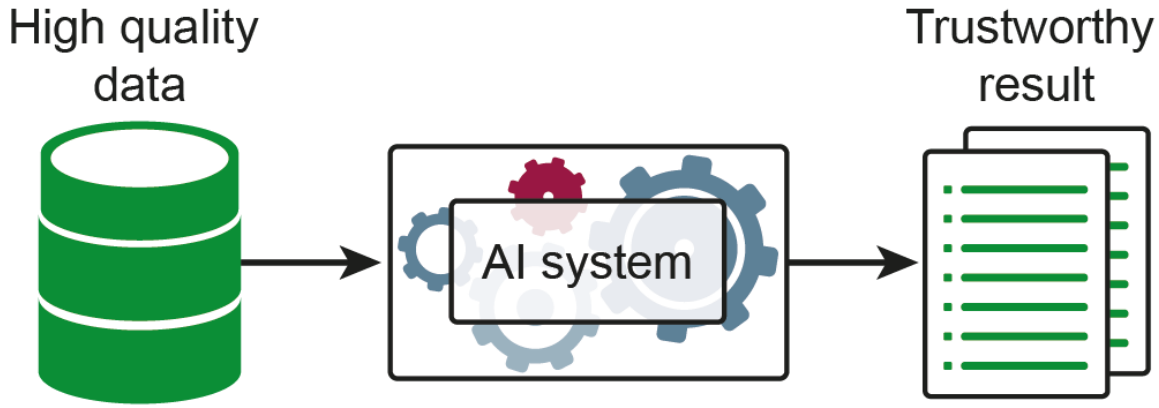
Standards for data quality:

→ **Content level** of the data and ***correctness***
→ **Traceability of data** (including data sources)
→ ***Completeness*** and **representativeness**
→ **Availability** and ***timeliness***

Motivate high-quality and secure **sensors**.

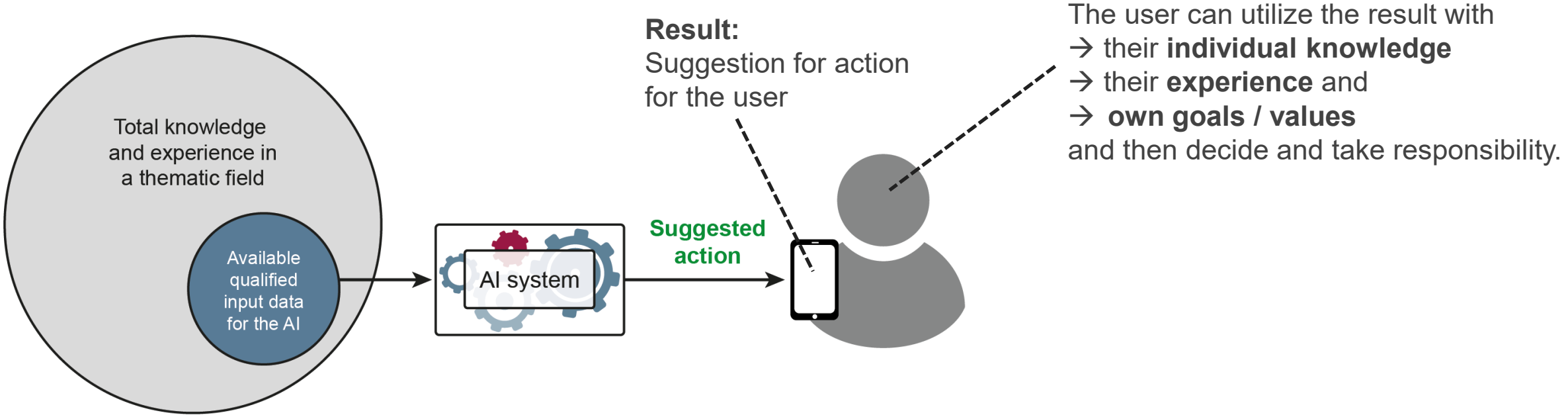
Further aspects for increasing quality:

→ Establish data pools
→ Promote the exchange of data
→ Create interoperability
→ Push open data strategy



Trustworthiness of result
→ Dealing with the AI result

- **„Keep the human in the loop“**
  - AI results are understood as a **suggestion for action** for the user.
  - *This concept promotes the user's **self-determination** and increases trust*

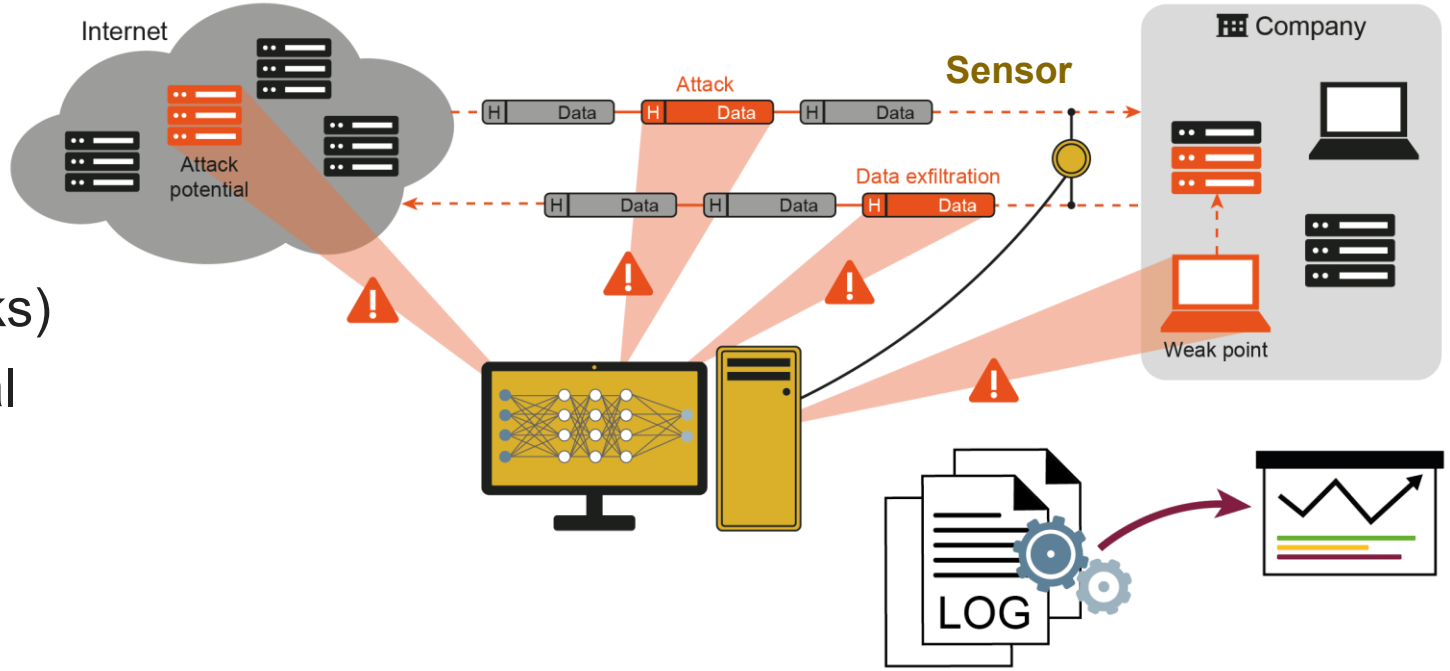


- **Automated applications such as autonomous driving**
  - ***High effort:*** *simulation, testing and validation*
  - ***Remaining risk:*** *responsibility, liability and insurance*

Artificial intelligence for IT security
→ Detecting attacks

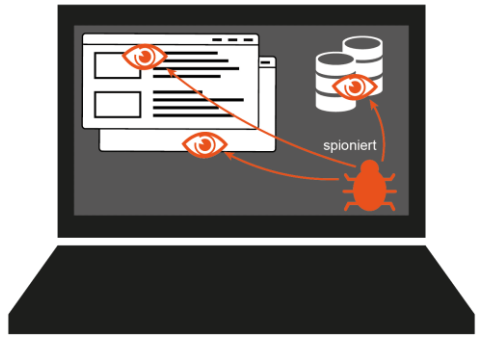
Increase the **detection rate** of **attacks**

- Network, IT end devices, servers, IoT devices, cloud applications ...
- Adaptive models (detect new attacks)
- Anomalies: normal versus abnormal



Other areas in which improved **detection** plays an important role:

- Networks / Internet
- *Malware*, spam …
- Fake news
- *Deepfake*
- ...



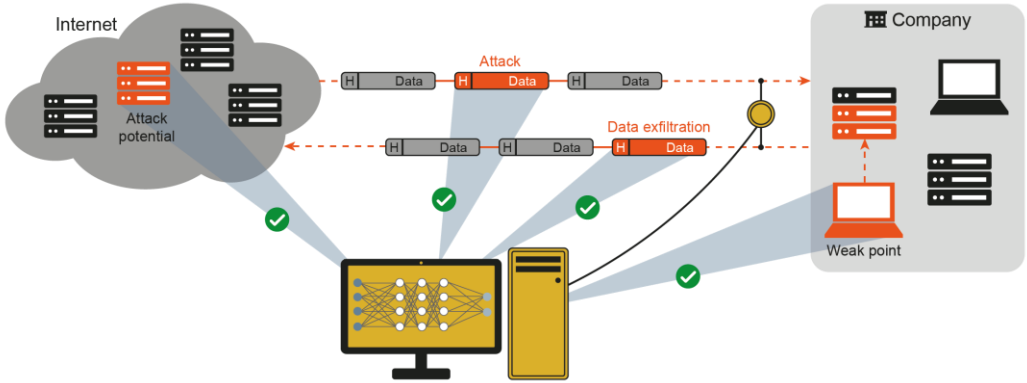
Attackers use artificial intelligence
→ Possible attacks

Identify attack vectors

- **Target**: Identify vulnerabilities / security gaps in the victim IT systems with the help of AI to assess **attack possibilities** / **potentials** and **carry out attacks**

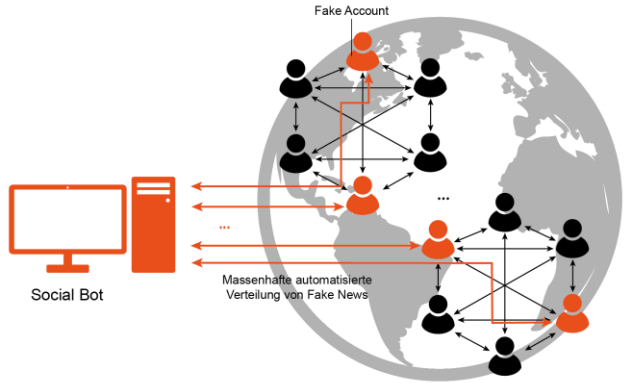
Overcoming IT security measures

- **Target:** “Light spots” (thresholds, sequences ...) of the anomaly detection system are **analyzed with the help of AI** to implement **undetected attacks**.



Automated attack

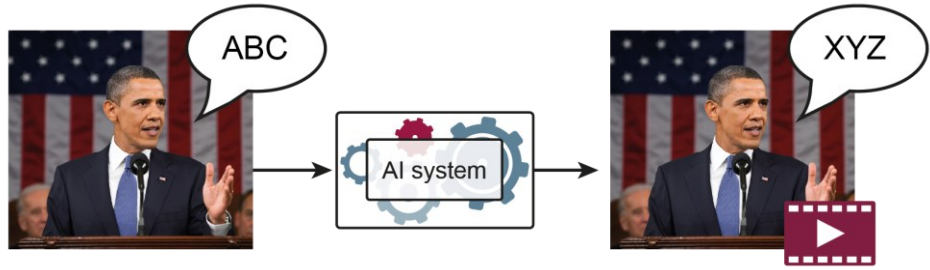
- **Target:** Social bots to **automatically generate fake accounts** and **fake news** to manipulate people.



Attackers use artificial intelligence
→ ChatBots like ChatGPT

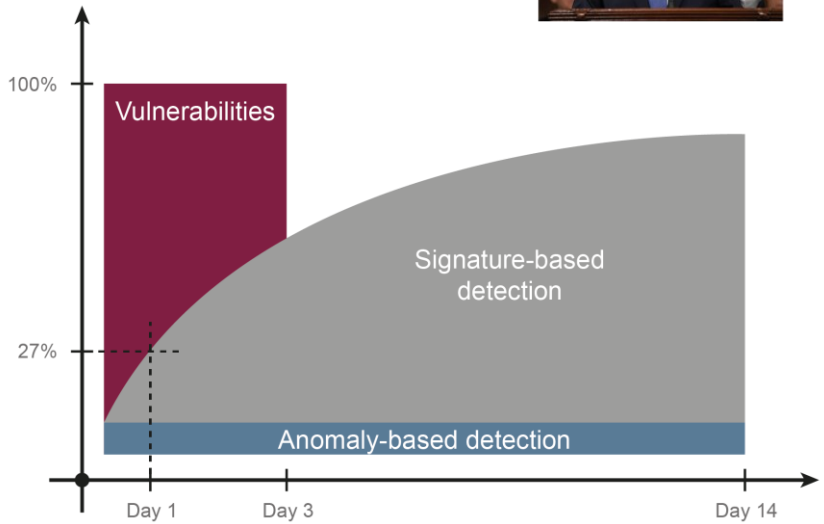
Social Engineering+++

- Spear phishing (first step for most attacks)
- Fake news (manipulation of people/societies)
- Deepfake (CEO fraud)



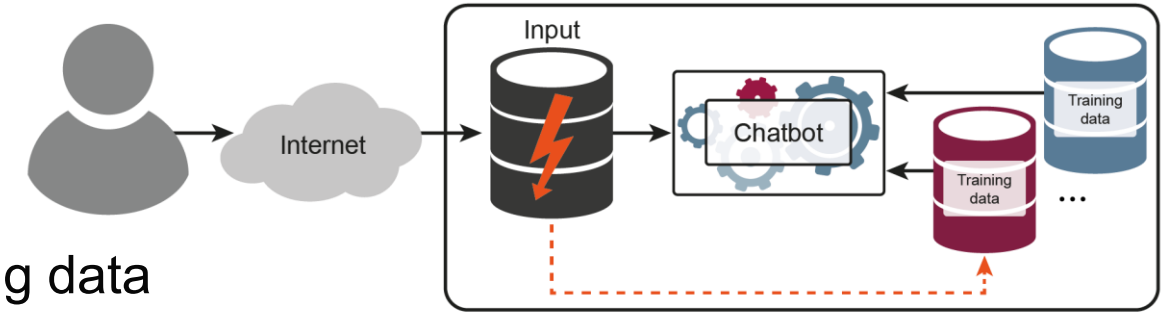
Faster development of attacks

- Polymorphic malware



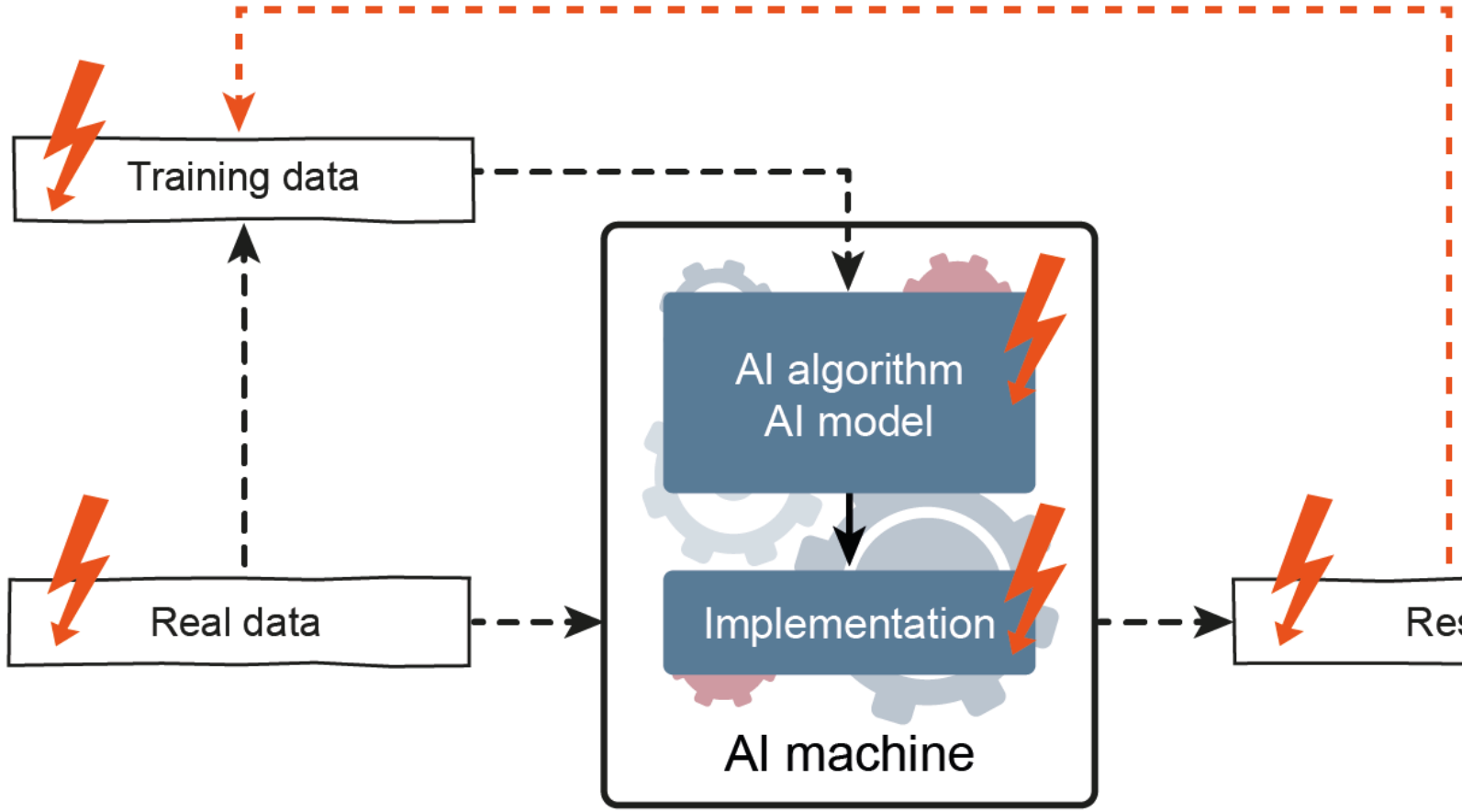
Confidential information as input to a central service

- Additional attack surface of confidential information
- Unwanted integration of confidential information into the training data



IT security for artificial intelligence

→ Attacks



Attacks on AI systems

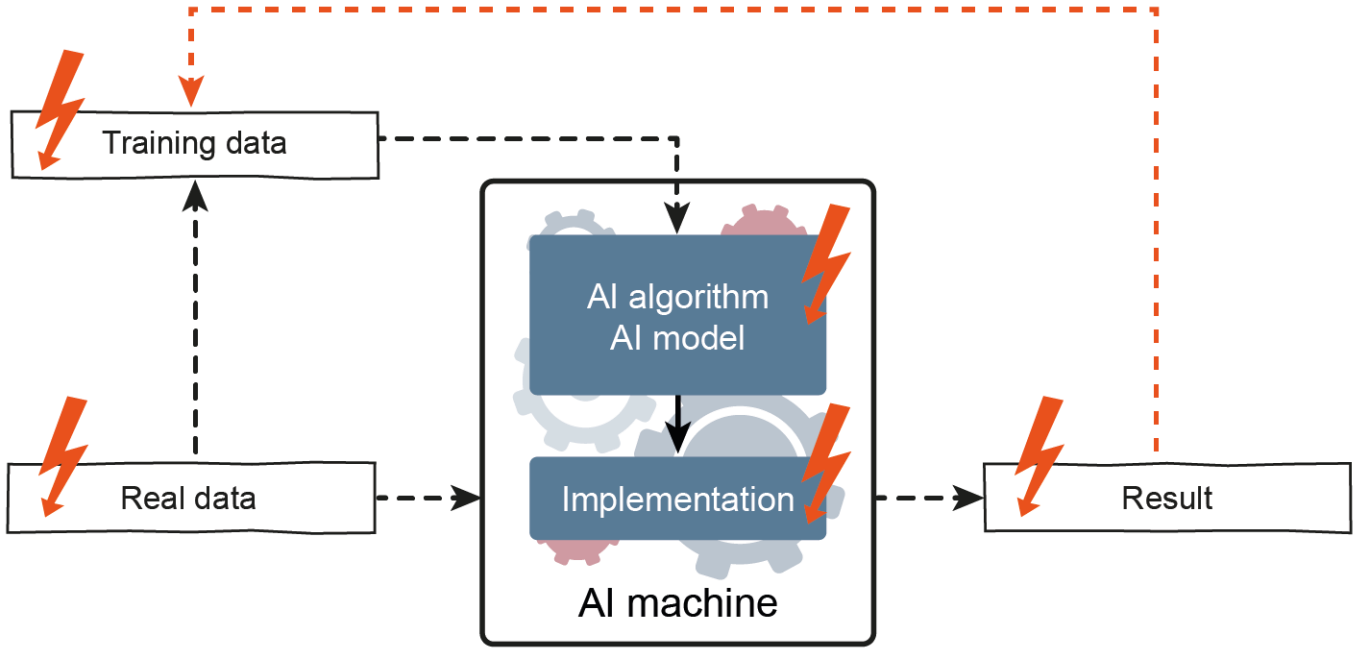
→ *Poisoning Attack*

→ *Evasion Attack*

→ *Exploratory Attack*

→ *…*

IT security for artificial intelligence
→ IT security measure



Use of high-quality AI technology

IT security goals:

→ ***Integrity***
(detection of data manipulation)

→ ***Confidentiality***
(protection of business secrets)

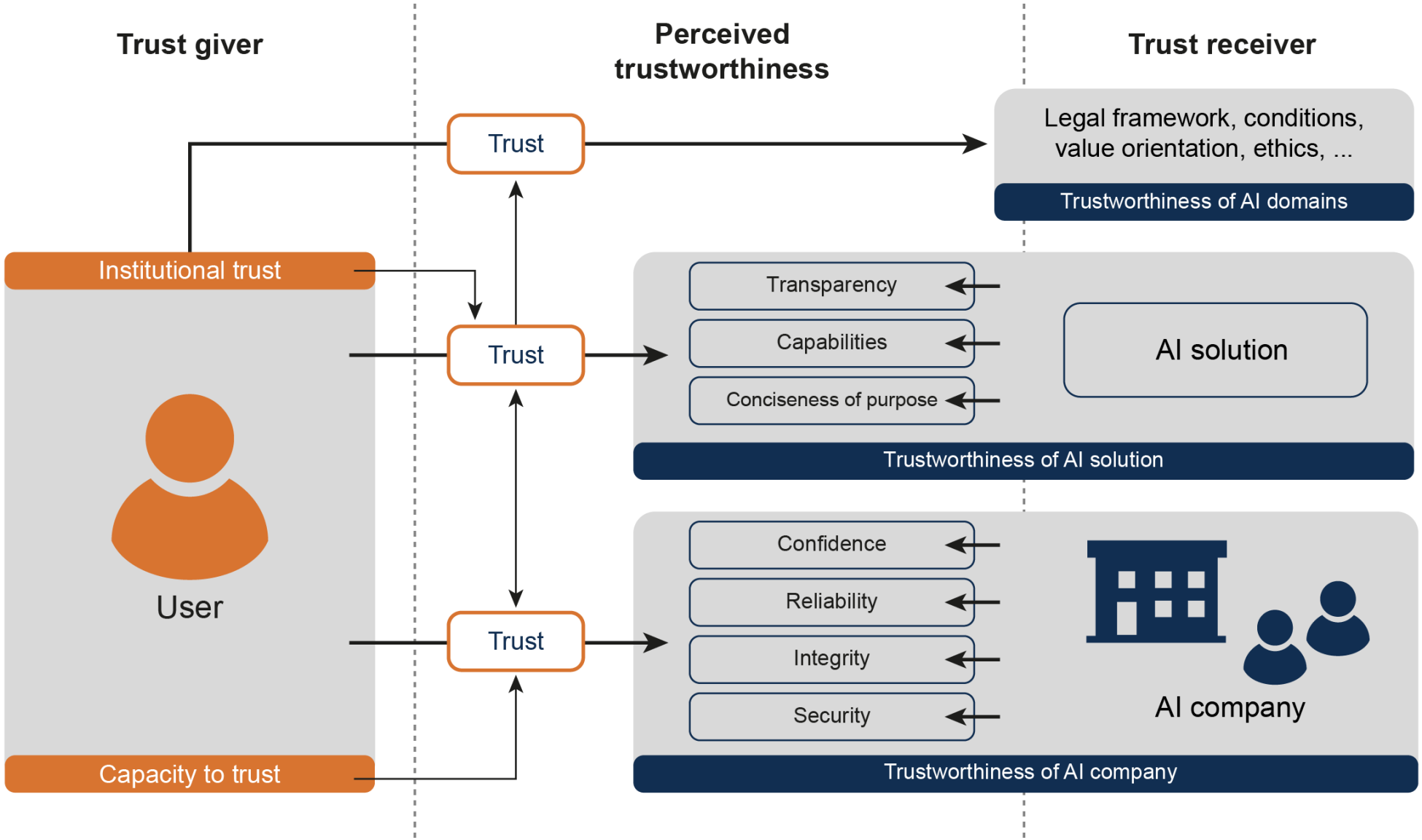
→ ***Data protection***
(protection of personal data)

→ ***Availability***
(of the application and results)

Use state-of-the-art technology for IT security measures

for protection of
→ the **data** (training data, real data/input data, results),
→ the **AI engine** (algorithms, models, implementation) and
→ the **application**

Artificial intelligence needs trust
→ Trustworthiness model



German-American IT Security Forum

IT Security Association Germany (TeleTrusT) in cooperation with GACC – German American Chamber of Commerce

San Francisco, 28.04.2025

AI and IT security
→ more security, more threats

We need AI in IT security,
but we also need to actively reduce the risks of AI!

Prof. Dr. Norbert Pohlmann

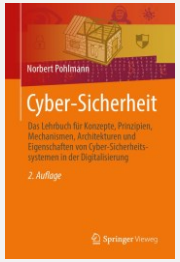
Professor for Cyber Security
Director of the Institute for Internet Security – if(is)
Chairman of the board of the IT Security Association TeleTrusT
Member of the board of the Internet industry association eco.

Anhang / Credits

Wir empfehlen

Cyber-Sicherheit

Das **Lehrbuch** für Konzepte, Mechanismen, Architekturen und Eigenschaften von Cyber-Sicherheitssystemen in der Digitalisierung“, Springer Vieweg Verlag, Wiesbaden 2022
https://norbert-pohlmann.com/cyber-sicherheit/



7. Sinn im Internet (Cyberschutzraum)

https://www.youtube.com/cyberschutzraum

CYBER SCHUTZRAUM

Master Internet-Sicherheit

https://it-sicherheit.de/master-studieren/

(M) MASTER internet-sicherheit.

Glossar Cyber-Sicherheit

https://norbert-pohlmann.com/category/glossar-cyber-sicherheit/

Cyber-Sicherheit GLOSSAR

It‘s all about Trust!

https://vertrauenswürdigkeit.com/

Besuchen und abonnieren Sie uns :-)

WWW
https://www.internet-sicherheit.de

Facebook
https://www.facebook.com/Internet.Sicherheit.ifis

Twitter
https://twitter.com/_ifis

https://twitter.com/ProfPohlmann

YouTube
https://www.youtube.com/user/InternetSicherheitDE/

Prof. Norbert Pohlmann
https://norbert-pohlmann.com/

Der Marktplatz IT-Sicherheit

(IT-Sicherheits-) Anbieter, Lösungen, Jobs, Veranstaltungen und Hilfestellungen (Ratgeber, IT-Sicherheitstipps, Glossar, u.v.m.) leicht & einfach finden.
https://www.it-sicherheit.de/

Quellen Bildmaterial

Eingebettete Piktogramme: Institut für Internet-Sicherheit – if(is)

Literatur

N. Pohlmann, S. Schmidt: „Der Virtuelle IT-Sicherheitsberater – Künstliche Intelligenz (KI) ergänzt statische Anomalien-Erkennung und signaturbasierte Intrusion Detection“, IT-Sicherheit – Management und Praxis, DATAKONTEXT-Fachverlag, 05/2009

D. Petersen, N. Pohlmann: "Ideales Internet-Frühwarnsystem", DuD Datenschutz und Datensicherheit – Recht und Sicherheit in Informationsverarbeitung und Kommunikation, Vieweg Verlag, 02/2011

M. Fourné, D. Petersen, N. Pohlmann: “Attack-Test and Verification Systems, Steps Towards Verifiable Anomaly Detection“. In Proceedings der INFORMATIK 2013 - Informatik angepasst an Mensch, Organisation und Umwelt, Hrsg.: Matthias Horbach, GI, Bonn 2013

U. Coester, N. Pohlmann: „Verlieren wir schleichend die Kontrolle über unser Handeln? Autonomie hat oberste Priorität“, BI-SPEKTRUM Fachzeitschrift für Business Intelligence und Data Warehousing, 05-2015

U. Coester, N. Pohlmann: „Ethik und künstliche Intelligenz – Wer macht die Spielregeln für die KI?“, IT & Production – Zeitschrift für erfolgreiche Produktion, TeDo Verlag, 2019

N. Pohlmann: „Künstliche Intelligenz und Cybersicherheit – Diskussionsgrundlage für den Digitalgipfel 2018“ https://norbert-pohlmann.com/app/uploads/2018/12/Künstliche-Intelligenz-und-Cybersicherheit-Diskussionsgrundlage-für-den-Digitalgipfel-2018-Prof.-Norbert-Pohlmann.pdf

N. Pohlmann: „Künstliche Intelligenz und Cybersicherheit - Unausgegoren aber notwendig“, IT-Sicherheit – Fachmagazin für Informationssicherheit und Compliance, DATAKONTEXT-Fachverlag, 1/2019

U. Coester, N. Pohlmann: „Wie können wir der KI vertrauen? - Mechanismus für gute Ergebnisse“, IT & Production – Zeitschrift für erfolgreiche Produktion, Technik-Dokumentations-Verlag, Ausgabe 2020/21

D. Adler, N. Demir, N. Pohlmann: „Angriffe auf die Künstliche Intelligenz – Bedrohungen und Schutzmaßnahmen“, IT-Sicherheit – Mittelstandsmagazin für Informationssicherheit und Datenschutz, DATAKONTEXT-Fachverlag, 1/2023

P. Farwick, Pohlmann: „Chancen und Risiken von ChatGPT – Vom angemessenen Umgang mit künstlicher Sprachintelligenz“, IT-Sicherheit – Mittelstandsmagazin für Informationssicherheit und Datenschutz, DATAKONTEXT-Fachverlag, 4/2023

N. Pohlmann: Lehrbuch „Cyber-Sicherheit“, Springer Vieweg Verlag, Wiesbaden 2022
Druckausgabe (ISBN 978-3-658-36242-3) und eBook (ISBN 978-3-658-36243-0).

Weitere Artikel siehe: https://norbert-pohlmann.com/artikel/