Westfälische Hochschule
Gelsenkirchen Bocholt Recklinghausen
University of Applied Sciences

Artificial Intelligence (AI) for Cyber Security

Prof. Dr. (TU NN)

Norbert Pohlmann

Professor** for **Cyber Security

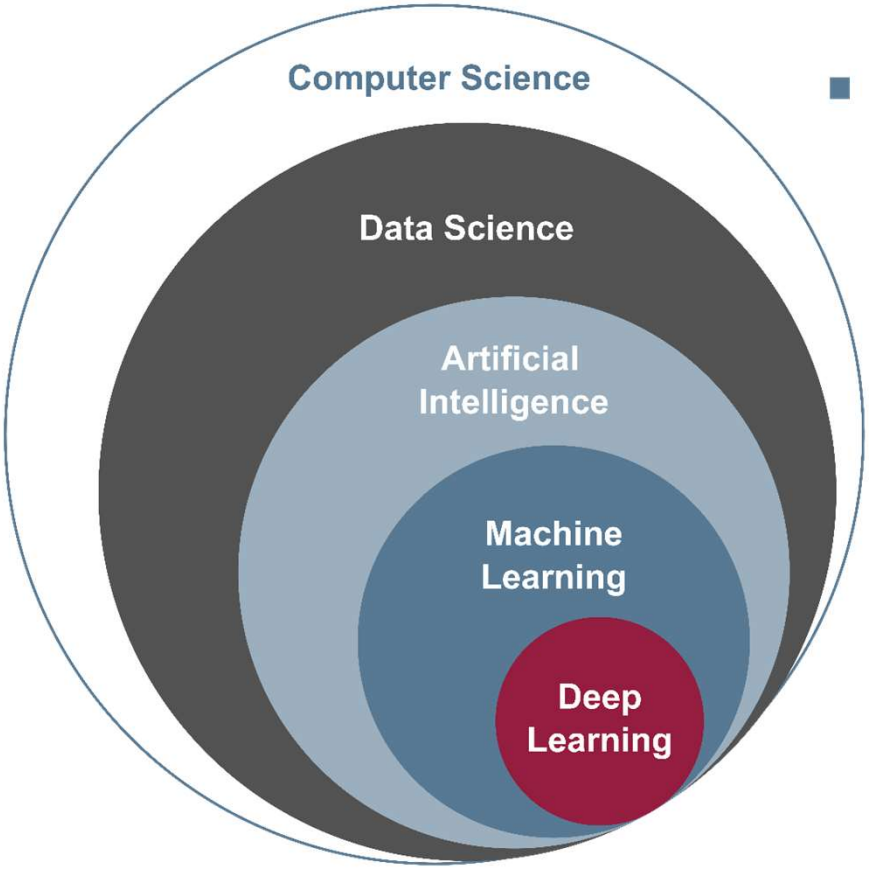
Director** of the **Institute for Internet Security – if(is)

Chairman of the board of the IT Security Association TeleTrusT

Member of the board** of the **Internet industry association eco.

if(is) internet security.

Classification
→ Artificial intelligence (AI)

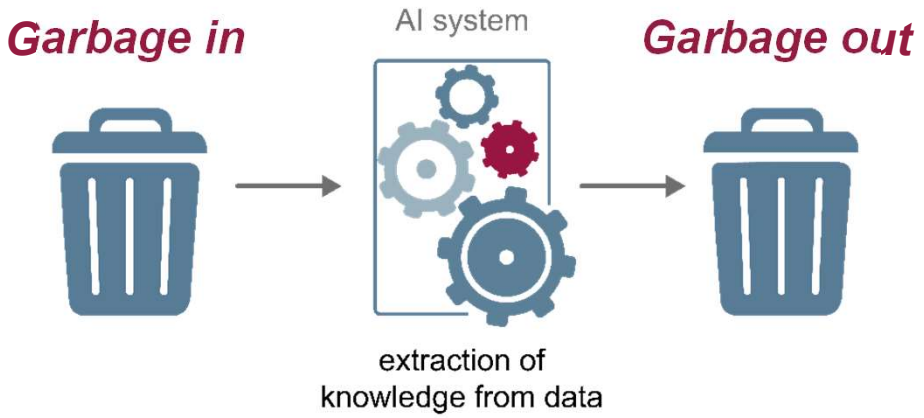


- **Data science** generally refers to the **extraction of knowledge from data.**
- **Artificial intelligence** translates intelligent behavior into algorithms.
  - **Strong "Artificial Intelligence"** *automatically replicate „human-like intelligence“.*
  - Superintelligence, **Singularity** *(“Machine” improves itself, is more intelligent than humans … **future**)*
    - **Weak “artificial intelligence”** *(machine learning – successfully implemented today)*
    - **Machine learning** is "artificial" **generation of knowledge from experience (in data)** by computer.
      - **Deep learning** is an important **improvement** of machine learning

Large Language Model (LLM) like ChatGPT

Trustworthiness of AI
→ Quality of the data

Paradigm



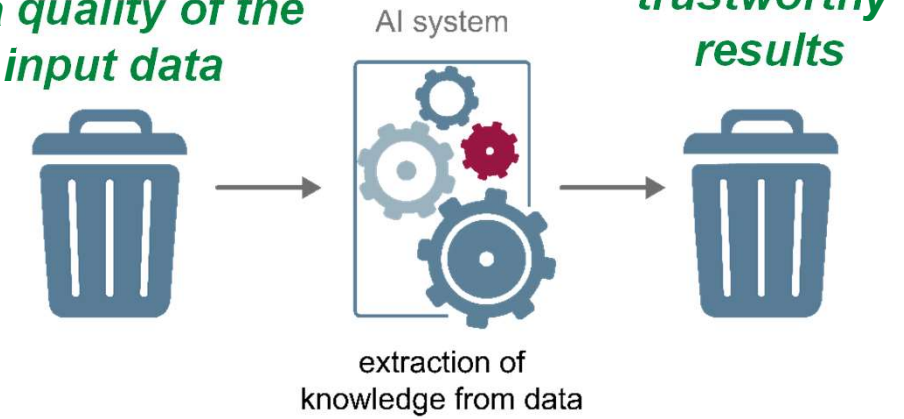
Standards for data quality:

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

Motivate high quality and secure **sensors**

high data quality of the input data

qualitative, trustworthy results

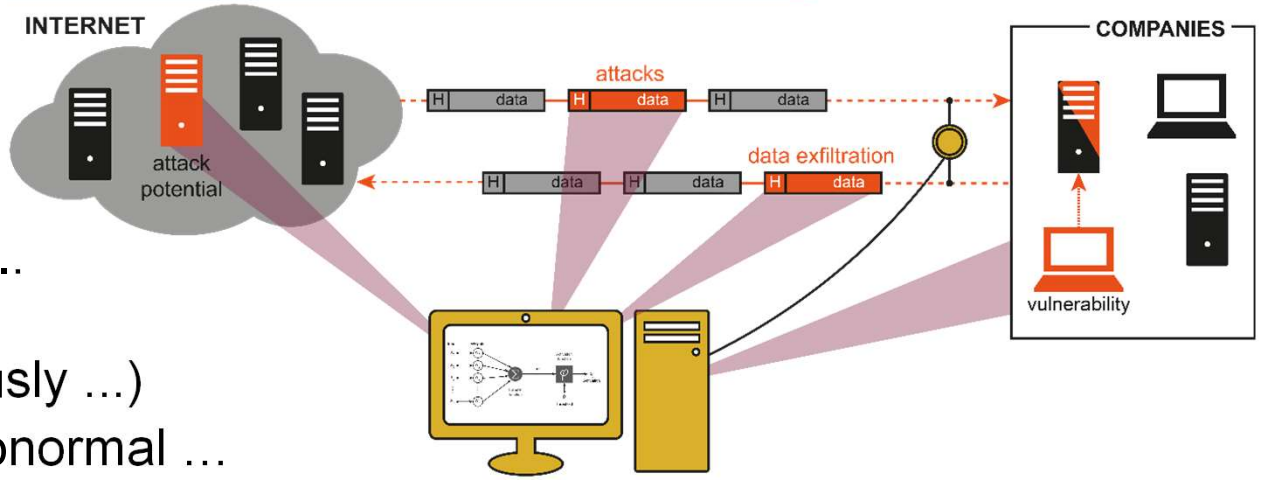


Other Ideas:

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

Artificial intelligence
→ for cyber security

- Increasing the **detection rate** of attacks
  - Network, IT end devices ...
  - adaptive models (independently, continuously ...)
  - Difference: normal and abnormal …



innovative detection of malicious network traffic

- **Support / Relief from cyber security experts** (of whom we do not have enough)
  - Finding **important** security-relevant events (prioritization)
  - **(Partial) autonomy** in response … resilience ...

- **Improvements** to existing **cyber security solutions**
  - AI contributes to increased impact and robustness
  - For example: risk-based and adaptive authentication

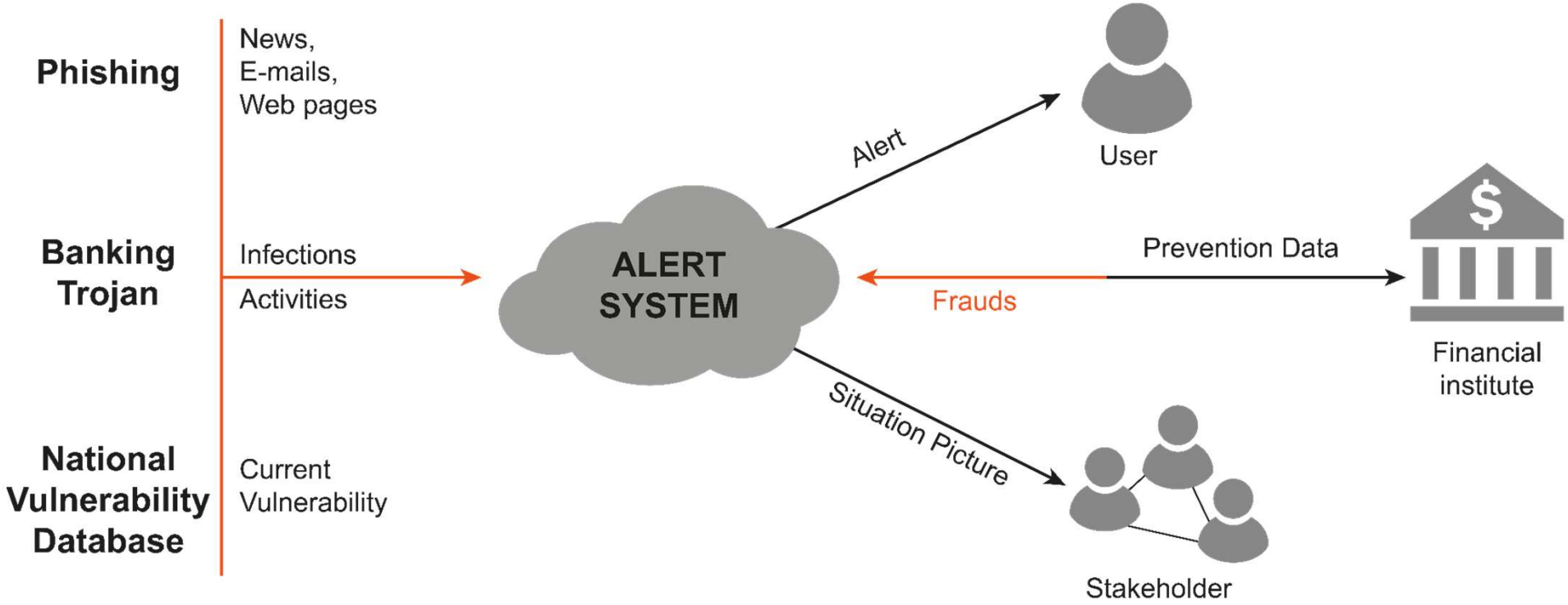
- **Further examples:** Detection from malware, spam, fake-news, deepfake ... secure software development, IT forensics, threat intelligence …

Research project
→ Alert-System for online banking

- **How could a solution look like?**
  - Warnings in the event of an increased risk situation (online banking)
    → enable the bank customer and the bank to react quickly and appropriately
  - Instruct the users when there are dangers
    → so that the bank customer can behave "correctly"

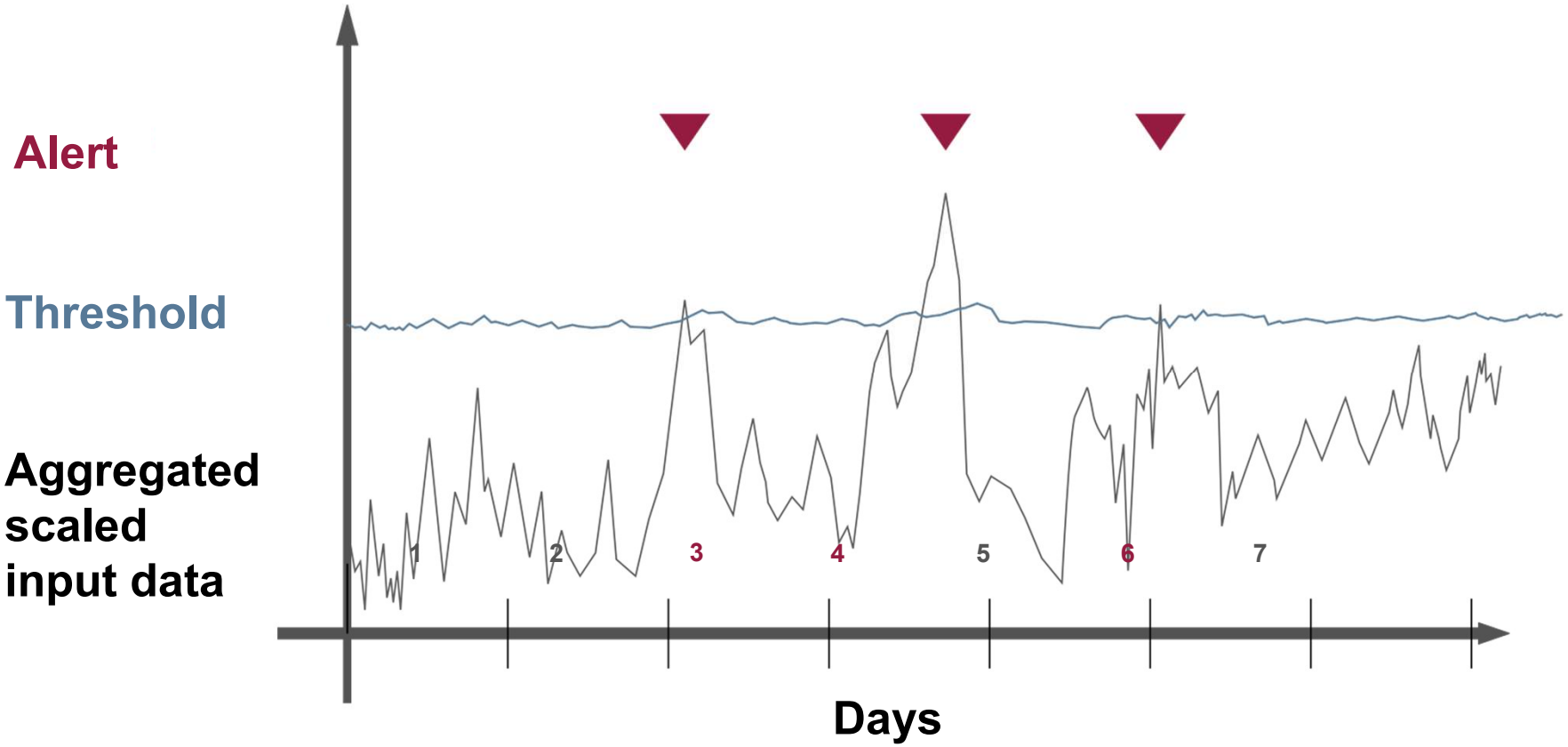
- **Approach of the alert system**
  - Identify **security metrics** for fraud
  - Determine **danger situation** with AI
  - **Warn** users and banks

Alert-System for online banking
→ Basic concept



- News (phishing attack) – we have received from the “Stackoverflow Network”
- **E-mail** (phishing attack) – are from the „Spam Archive“
- Phishing **websites** – we have received from the „PhishTank“
- **Information** of banking Trojans (malware) - we got from anti-malware companies
- Relevant and current **vulnerabilities** we have retrieved from the NVD
- Successful **fraud cases** in online banking – were provides by the banking group

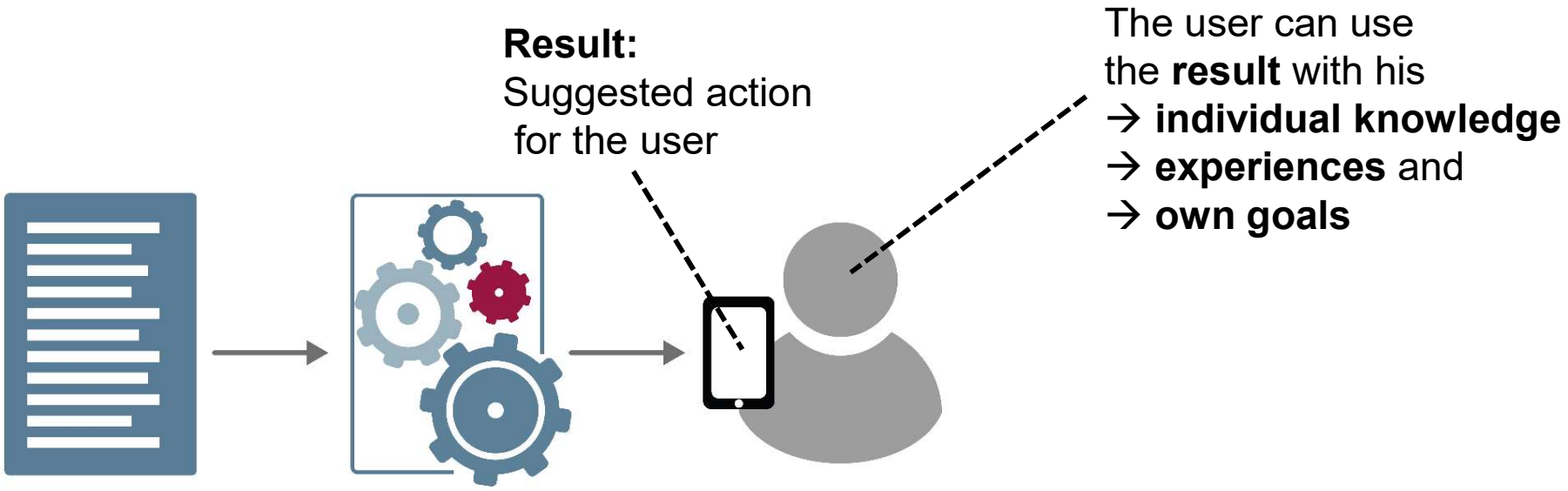
Alert-System for online banking
→ Result



- **Output:**
  - Predicted threat values on days 3, 4, and 6 exceed the threshold set for this alert system
  - because the threshold has been exceeded, an alert is triggered

Trustworthiness
→ Types of validation of results

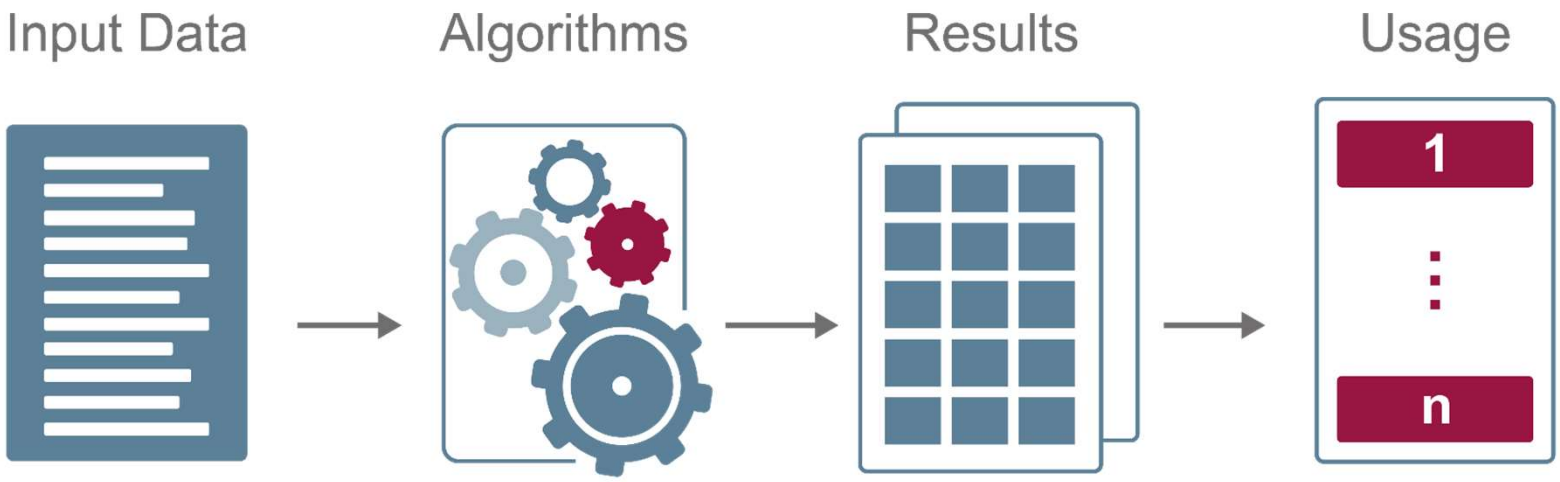
- **„Keep the human in the loop“**
  - AI result must be understood as a **recommendation for the** user.
  - This promotes the **self-determination** of users and increases their trustworthiness.



- **Automated applications** (e.g., autonomous driving)
  - Simulation, test and **validation**
  - Responsibility, **liability** and insurance

Attacks
→ on machine learning (AI)

Hackers attack and manipulate the workflow (“result”)

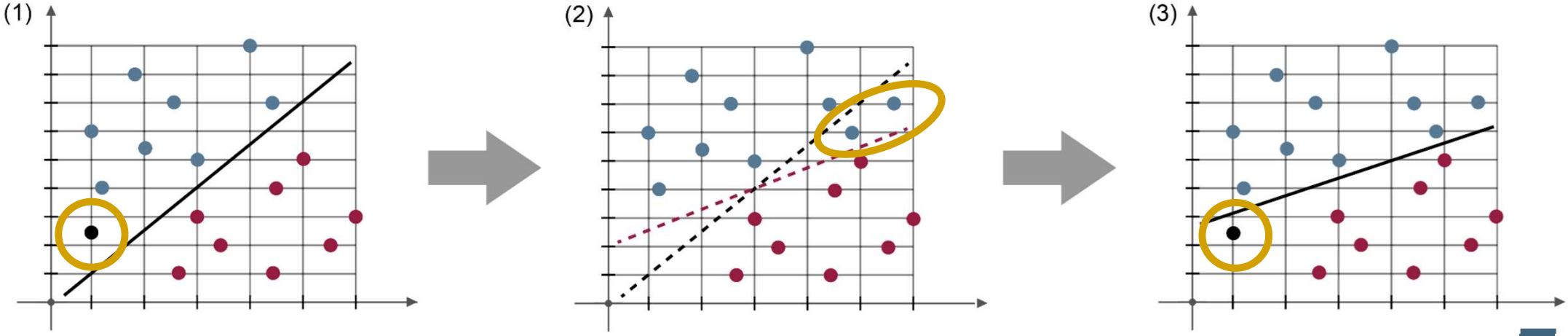


- ***Input data (input)***
- ***Algorithms / Models***
- ***Results (output)***
- ***Usage***

Attacks on machine learning

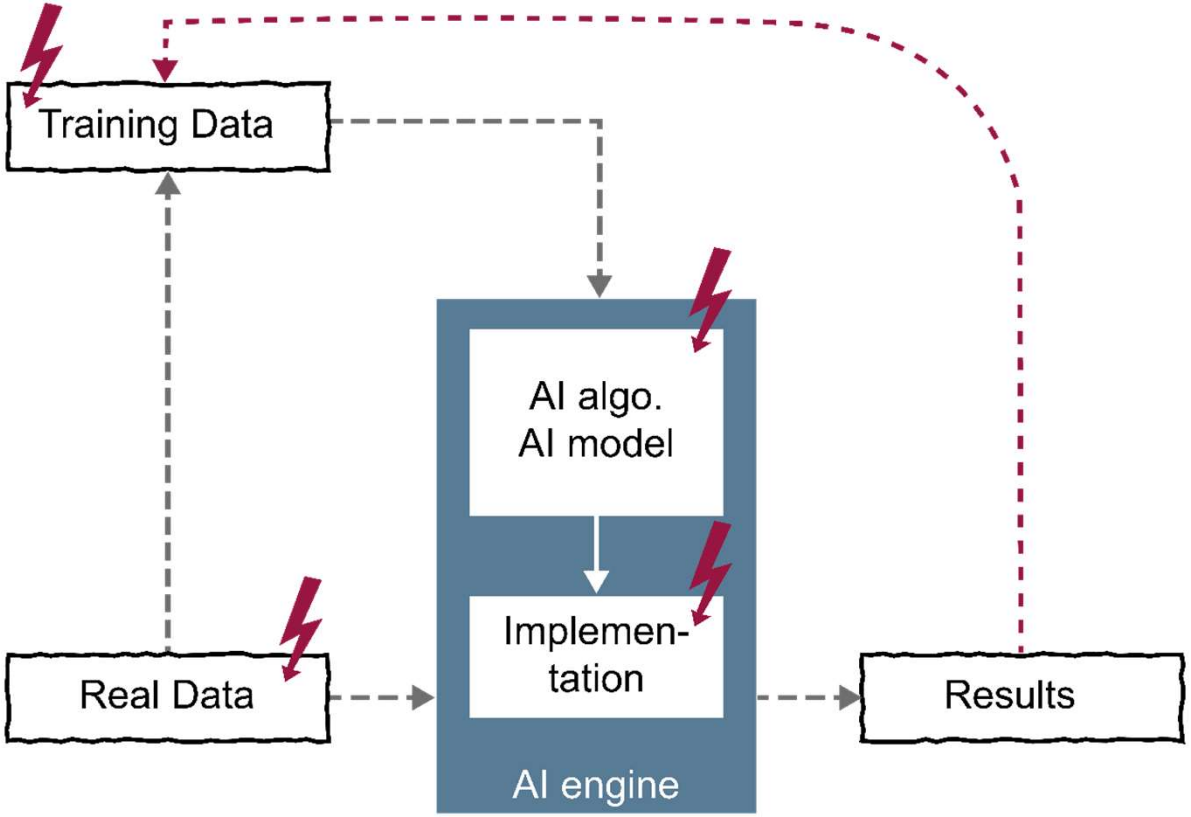
→ **Manipulation of training data** (Poisoning Attack)

- (1) **Normal classification** of a new input.
  (new black dot belongs to the blue class)
- **(2) Example: manipulation of training data**
  - Incorrectly classified data will be injected into the training phase as an attack (two more blue dots).
  - This manipulates the straight line of the model for classification (straight line becomes flatter).
- (3) This can be used by an attacker to create **wrong classifications**.
  (now the new black dot belongs to the red class)



Secure AI
→ Protection of the implementation and data

State of the art
cyber security measures
for protection
→ *the **data** (training, real, result),*
→ *the **AI engine** and*
→ *the **application***



Use of a
high quality
AI technology

Security goals:
→ ***Integrity***
(detection of data manipulation)
→ ***Confidentiality***
(protection of business secrets)
→ ***Data protection***
(protection of personal data)
→ ***Availability***
(of the application and results)

AI for Cyber Security
→ Result and outlook

- **AI / ML is an important technology in the flied of cyber security**
  - Detect threats, vulnerabilities, attacks ...
  - Support of cyber security experts
  - Secure software development
  - …
- **We need to secure our AI to be able to produce trustworthy results**
  - Hackers attack and manipulate data, algorithm/models and results
  - …
- **Balance of power for the future between attacker and defender**
  - The **attackers** use AI for their attacks **very successfully**
  - The defenders should do this more and also together
  - …

Westfälische Hochschule
Gelsenkirchen Bocholt Recklinghausen
University of Applied Sciences

Artificial Intelligence (AI) for Cyber Security

*With **Artificial Intelligence***
into a more secure digital future!

Prof. Dr. (TU NN)

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.

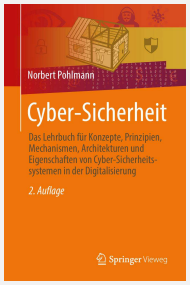
if(is) internet security.

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)

Literature

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

U. Coester, N. Pohlmann: „Diskriminierung und weniger Selbstbestimmung? Die Schattenseiten der Algorithmen“, tec4u, 12/17

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).

See more articles: https://norbert-pohlmann.com/artikel/