



**Westfälische  
Hochschule**

Gelsenkirchen Bocholt Recklinghausen  
University of Applied Sciences

# Künstliche Intelligenz (KI) *und* Cyber-Sicherheit

- **Cyber-Sicherheit** braucht **Künstliche Intelligenz**
- **Künstliche Intelligenz** braucht **Cyber-Sicherheit**

Prof. Dr. (TU NN)

**Norbert Pohlmann**

*Professor für Informationssicherheit*

*Leiter des Instituts für Internet-Sicherheit - if(is)*

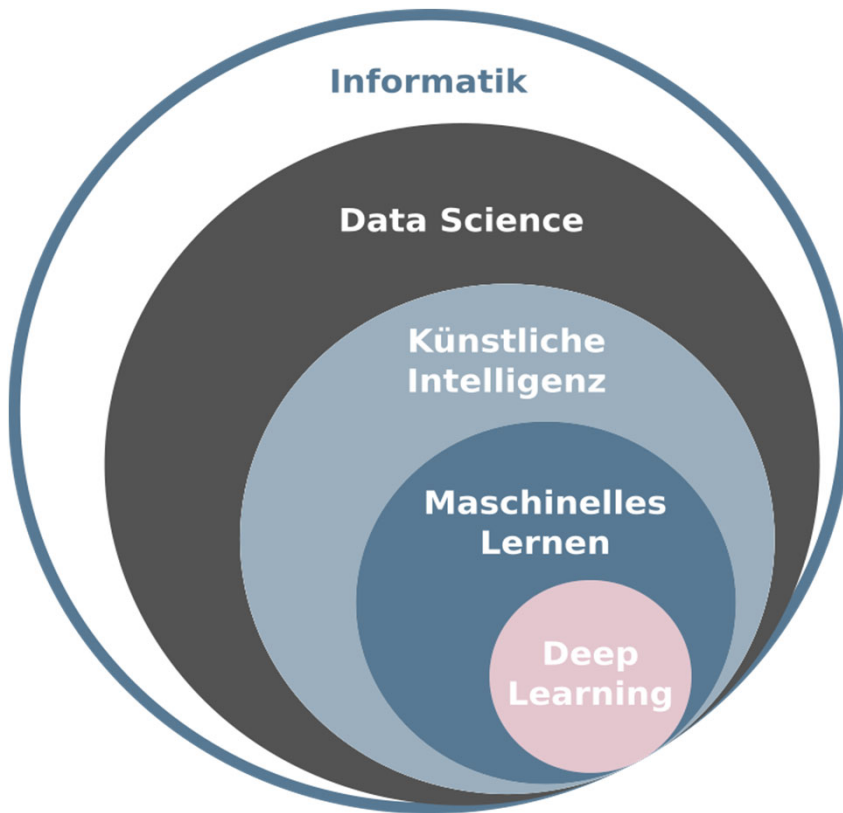
*Vorstandsmitglied des Verbands der Internetwirtschaft - eco*

*Vorstandsvorsitzender des Bundesverbands IT-Sicherheit - TeleTrust*

**if(is)**  
internet-sicherheit.

# Einordnung

→ (Künstliche Intelligenz) **Maschinelles Lernen**



- **Data Science** bezeichnet generell die **Extraktion von Wissen** aus Daten.
- **Starke „Künstliche Intelligenz“ (Zukunft)** soll automatisiert „**menschenähnliche Intelligenz**“ nachbilden. **Singularität** („Maschinen“ verbessern sich selbst, sind intelligenter als Menschen)
- **Schwache „Künstliche Intelligenz“ (heute)** **Maschinelles Lernen** ist ein Begriff für die „künstliche“ **Generierung von Wissen aus Erfahrung** (in Daten) durch Computer.

# Erfolgsfaktoren

→ (Künstliche Intelligenz) **Maschinelles Lernen**

## Leistungsfähigkeit der IT-Systeme

- enorme Steigerung (CPU, RAM, ..) der IT-Systeme, sehr viele CPU Kerne, ...
- Spezial-Hardware: GPUs, FPGA, TensorFlow PU (TPU), ...
- Frameworks, Cloud-Lösungen, ...

## Algorithmen

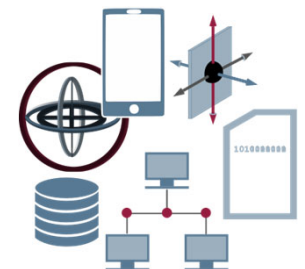
- Immer **bessere Algorithmen** (oftmals als OpenSource)
- Immer **mehr Erfahrung** mit dem Einsatz
- Immer **einfacherer Zugang** zu den Technologien und Diensten



## Daten

Immer **mehr** vorhandene (sicherheitsrelevante) **Daten** für die

- **Verteidiger**  
aber auch für die
- **Angreifer**



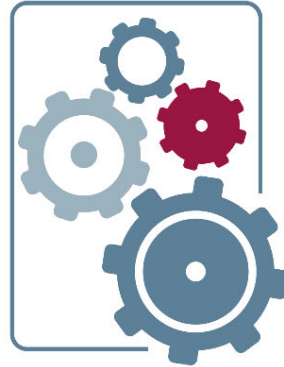
# Maschinelles Lernen

## → Workflow

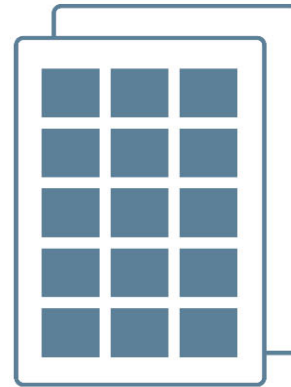
Eingabedaten



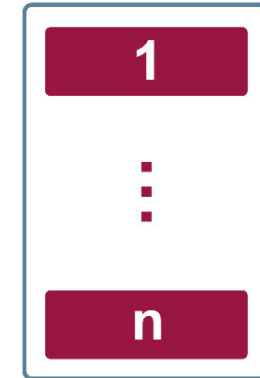
Algorithmus



Ergebnisse



Verwendung



### Eingangsdaten

Qualität: Inhalt, Vollständigkeiten, Repräsentativität, ... Aufbereitung

### Algorithmen (ML)

Support-Vector-Machine (SVM), k-Nearest-Neighbor (kNN), ... Deep Learning

### Ergebnisse

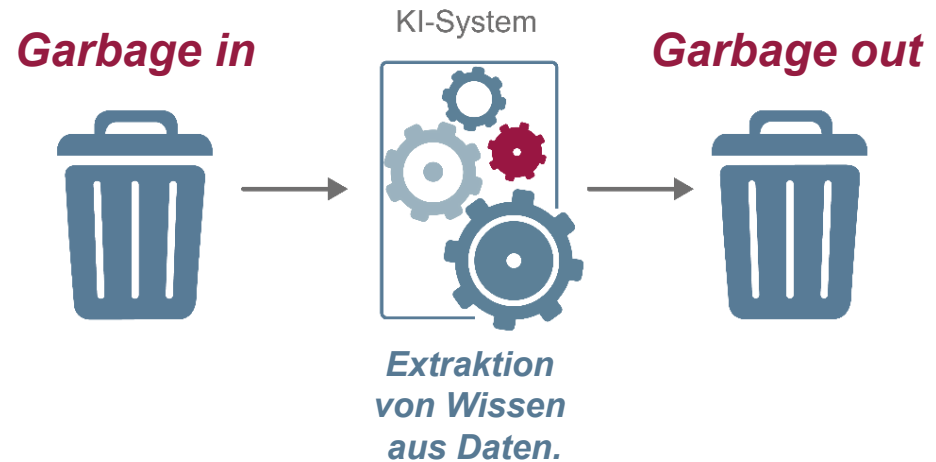
Ergebnisse (Wissen) aus der Verarbeitung (Algorithmus) der Eingangsdaten ...

### Verwendung

Die Anwendung entscheidet, wie Ergebnisse verwendet werden (*Vertrauen*).

# Vertrauenswürdigkeit → Qualität der Daten

## Paradigma



### Standards für die Datenqualität:

- Inalthöhe der Daten und **Korrektheit**
- **Nachvollziehbarkeit** (Datenquellen)
- Vollständigkeit und **Repräsentativität**
- Verfügbarkeit und Aktualität

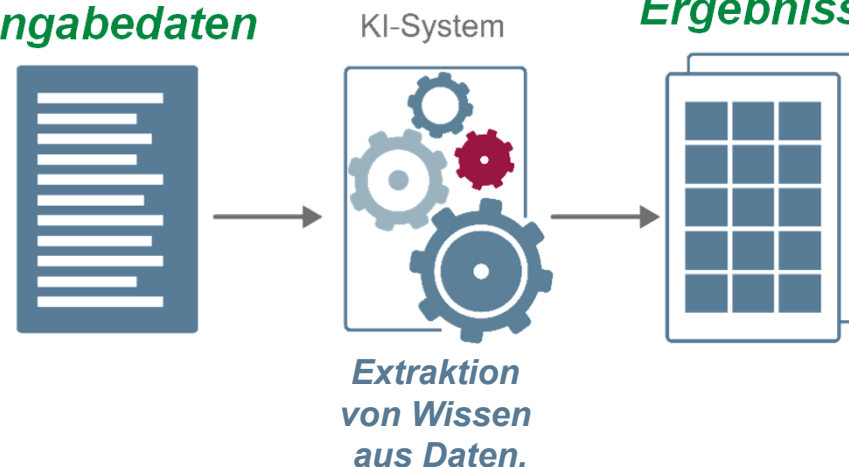
Qualitativ hochwertige und sichere Sensoren motivieren

*hohe Datenqualität der Eingabedaten*

*qualitative, vertrauenswürdige Ergebnisse*

### Weitere Aspekte zur Erhöhung der Qualität:

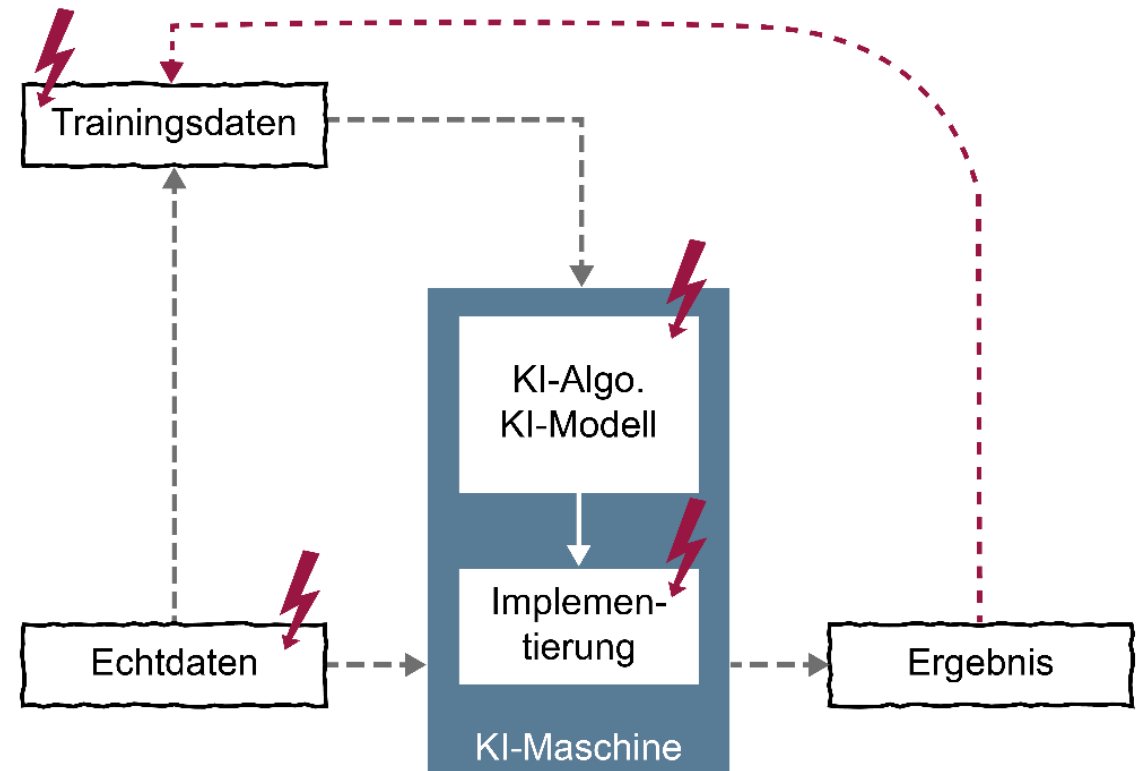
- Datenpools etablieren
- Austausch von Daten fördern
- Interoperabilität schaffen
- Open Data-Strategie puschen



# Risiken der KI aktiv reduzieren → Cyber-Sicherheitsmaßnahmen

## Stand der Technik an Cyber-Sicherheitsmaßnahmen zum Schutz

- der **Daten** (Training, Echt, Ergebnis),
- der **KI-Maschine** und
- der **Anwendung**



**Nutzung einer qualitativ hochwertigen  
KI-Technologie**

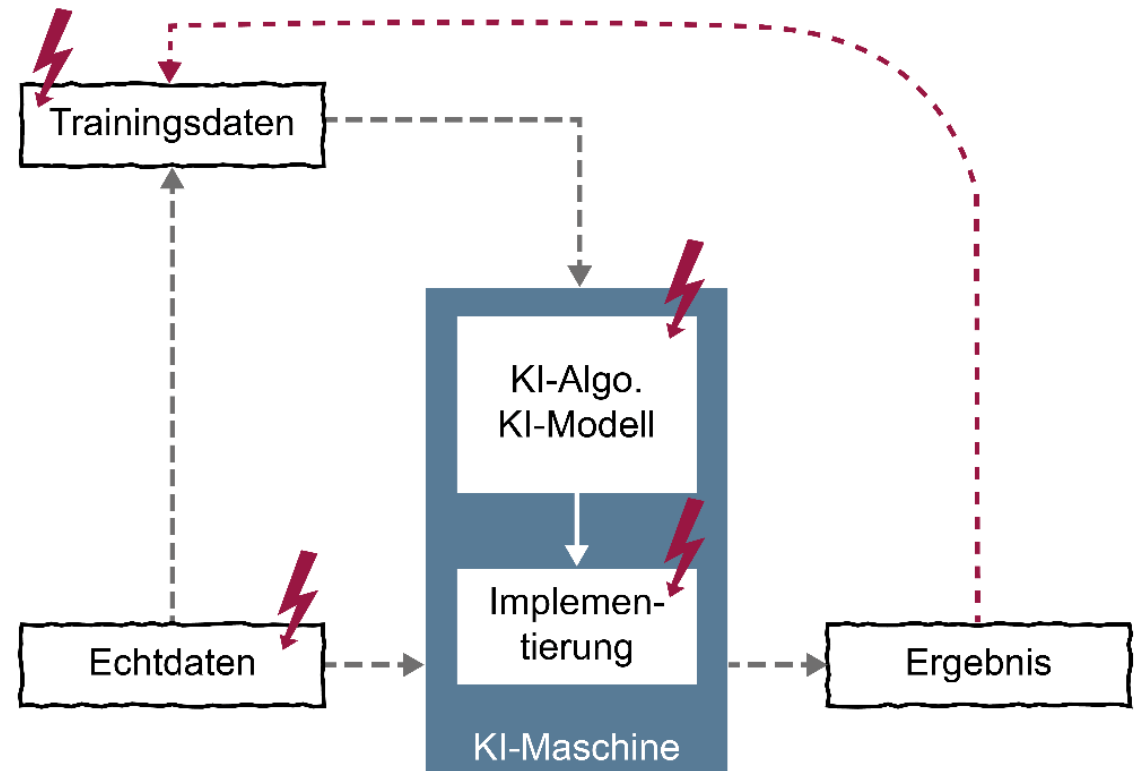
**(Evaluierung / Zertifizierung / Souveränität / GAIA-X)**

# Risiken der KI aktiv reduzieren

## → Schutzziele

### Schutzziele:

- **Integrität**  
(Erkennen von Manipulation der Daten)
- **Vertraulichkeit**  
(Wahrung von Geschäftsgeheimnissen)
- **Datenschutz**  
(Schutz von personenbezogenen Daten)
- **Verfügbarkeit**  
(der Anwendung und Ergebnisse)



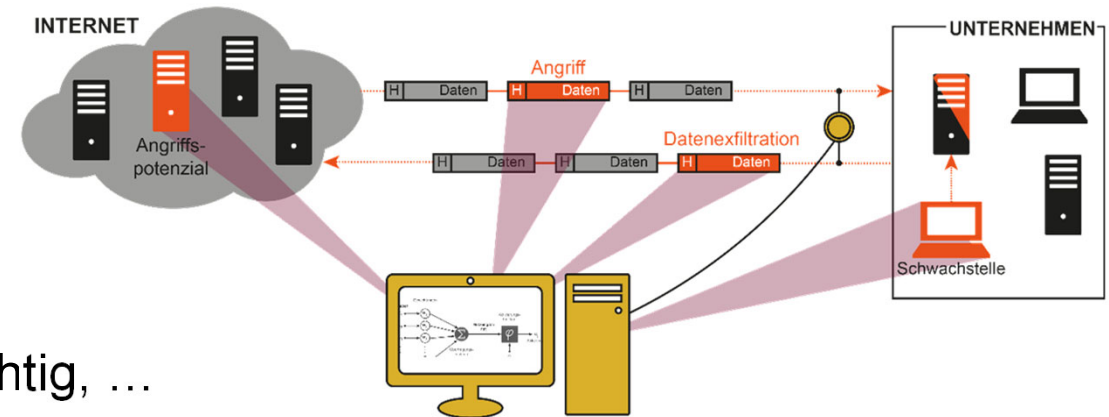
**Zusammenarbeit von erfahrenen  
KI- und Cyber-Sicherheitsexperten**  
(Aufbau / Sicherstellung vom Kompetenzen  
- Ergebnisse, Ethik, ...)

# Cyber-Sicherheit *braucht*

## → Künstliche Intelligenz - Übersicht

- Erhöhung der **Erkennungsrate** von **Angriffen**

- Netzwerk, IT-Endgeräte, ...
- adaptive Modelle
- Unterschied: normal und verdächtig, ...



- **Unterstützung / Entlastung** von **Cyber-Sicherheitsexperten**

- Erkennen von **wichtigen** sicherheitsrelevanten Ereignissen (*Priorisierung*)
- **(Teil-)Autonomie** bei Reaktionen, ... Erhöhung der Resilienz, ...

- **Verbesserungen** von bestehenden **Cyber-Sicherheitslösungen** durch **KI**

- zur Erhöhung der **Wirkung** und **Robustheit**,  
z.B.: Risiko-basierte und adaptive Authentifizierung
- um **Schäden** zu **vermeiden** und **Risiken** zu **minimieren!**



**Weitere Bereiche:** Erkennung von Malware, Spam, Fake-News, Deep-Fake, usw.  
sichere Softwareentwicklung, IT-Forensik, Threat Intelligence, ...



# Anwendungen von KI und CS (1/2)

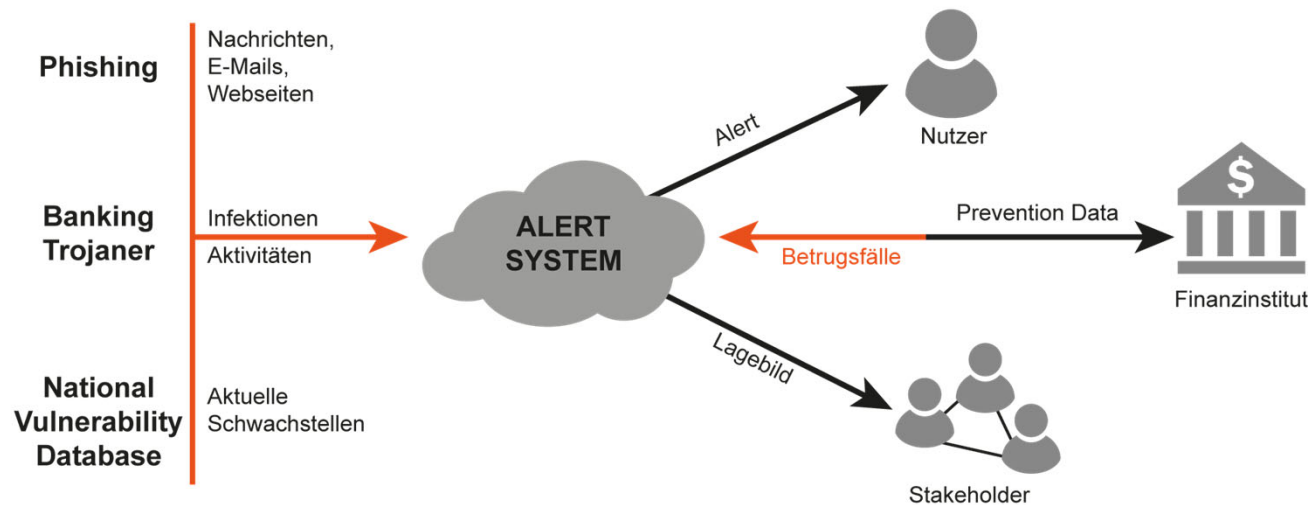
## → Alert-System für Online-Banking

- Wie könnte eine Lösung aussehen?
  - Tagesaktuelle Warnungen bei erhöhter Gefahrenlage (Online-Banking)  
→ **damit der Bankkunde und die Bank reagieren können**
  - Aufklärung der Nutzer, wenn Gefahren vorliegen  
→ **damit der Bankkunde sich „richtig“ verhalten kann**
- **Ansatz des Alert-Systems**
  - ***Sicherheitskennzahlen*** zum Betrug identifizieren
  - Mittels KI ***Gefahrenlage bestimmen***
  - Nutzer und Bank ***warnen***

# Alert-System für Online-Banking

→ Zahlen für den Testzeitraum von 456 Tage

- 1.904 Nachrichten (Phishing-Angriff) – „Stackoverflow-Netzwerk“
- 5.589 **E-Mail** (Phishing-Angriff) – „Spam Archive“
- 2.776 Phishing-**Webseiten** – „PhishTank“
- 23.184 **Infektionen** von Banking-Trojaner (Malware) – Anti-Malwarehersteller
- 875 relevante **Schwachstellen** (NVD)
- 459 erfolgreiche **Betrugsfälle** im Online-Banking - Bankengruppe



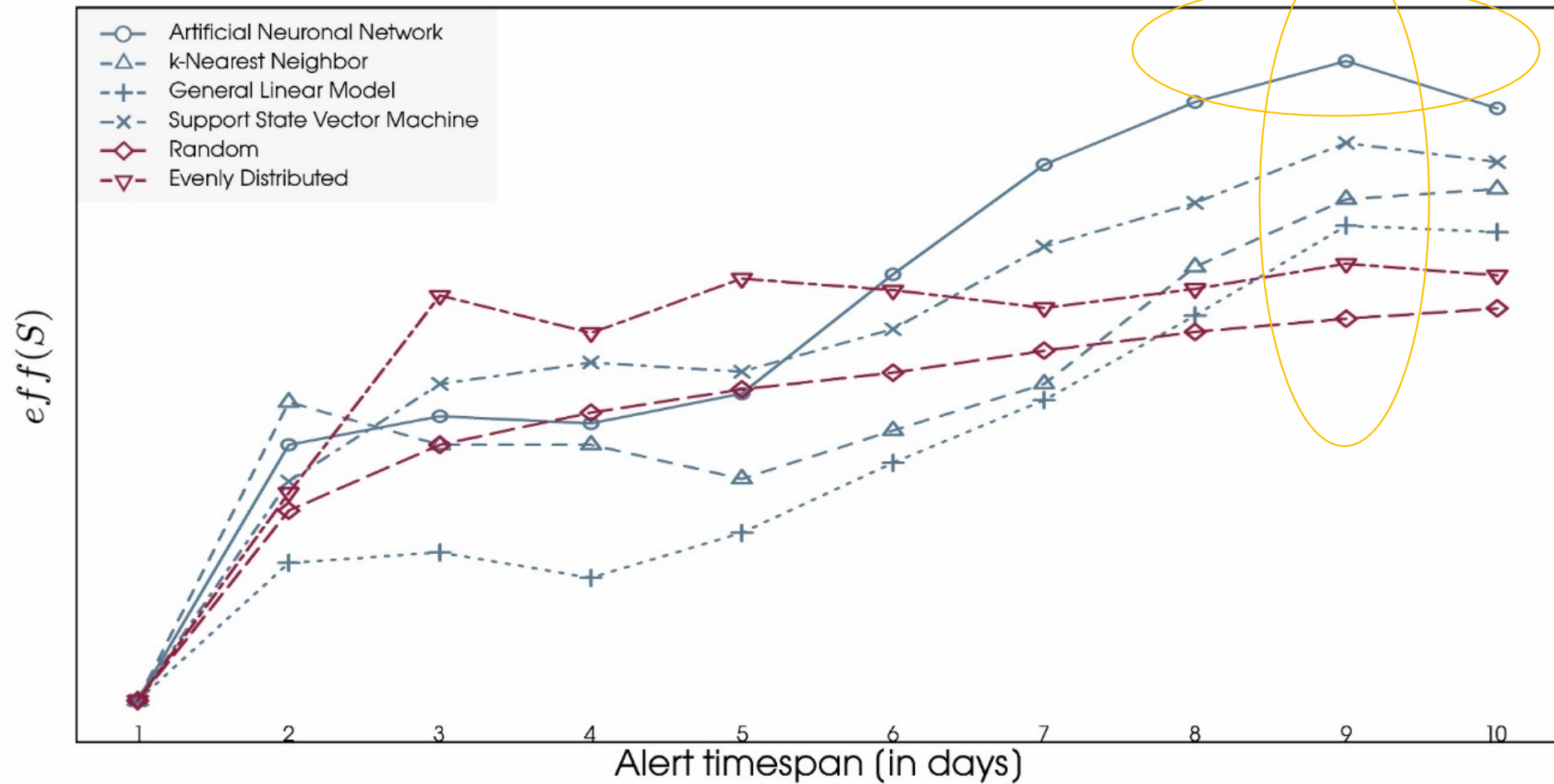
1/3 des Zeitraums zum Training (152 Tage) 2/3 zur Evaluation (304 Tage)

# Ergebnisse

## → Vergleich der verschiedenen Verfahren

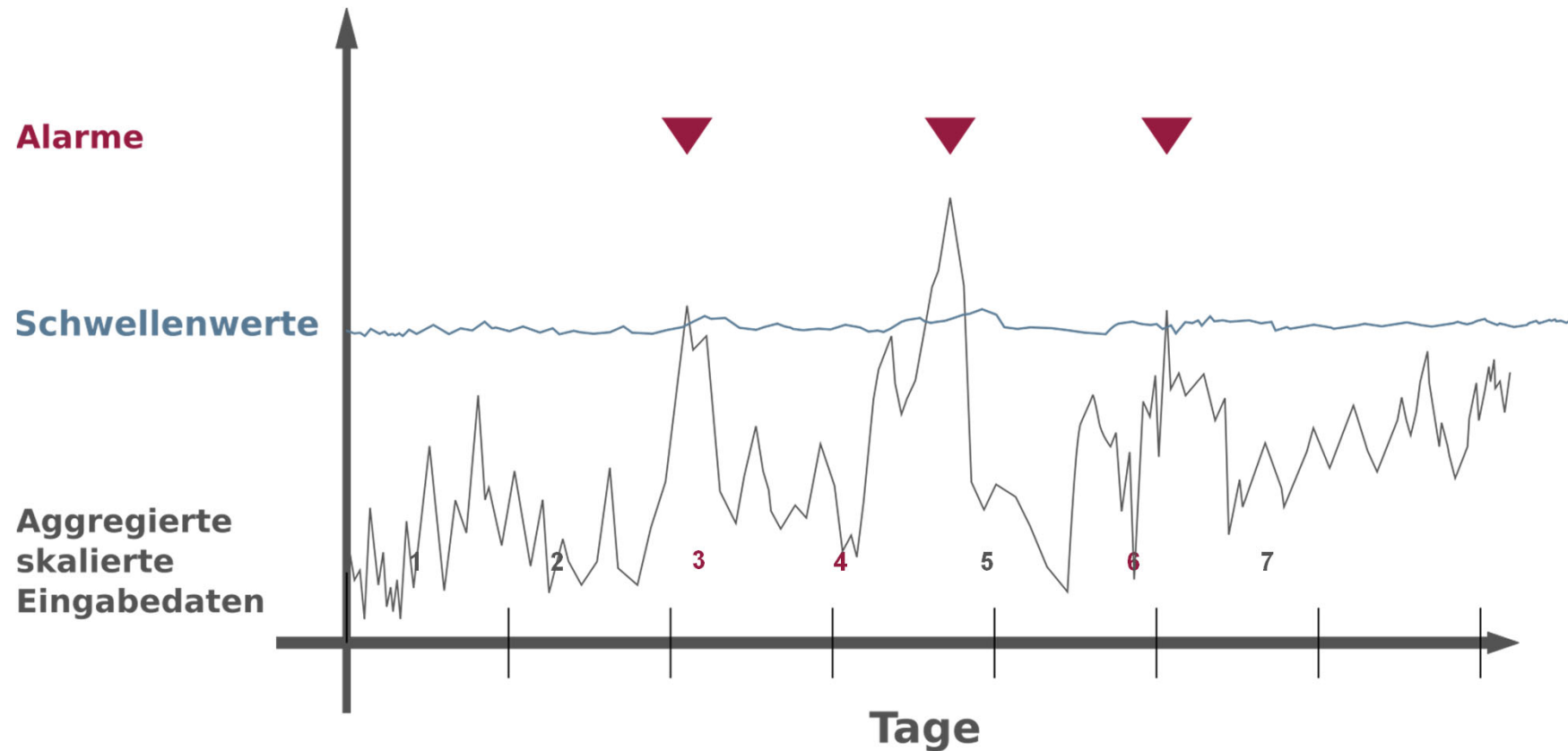
„Aber, drei Mal soviel Zeit für das Trainieren“

### Comparison of the different approaches



# Alert-System für Online-Banking

## → Ergebnis



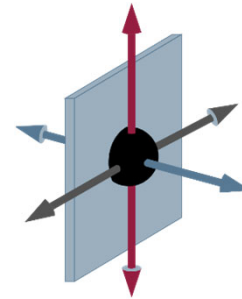
### ■ Output:

- Vorhergesagte Bedrohungswerte überschreiten an den Tagen 3, 4 und 6 den für dieses Alert-System eingestellten Schwellenwert.
- Da der Schwellenwert überschritten wurde, wird ein Alarm ausgelöst.

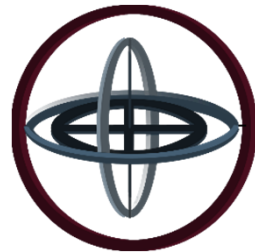
## → Passive Authentifikation

- Ein Nutzer wird automatisiert an der Art und Weise der Nutzung beim QR-Code Scannen erkannt.
- Während des gesamten Vorgangs werden passive biometrische Bewegungsdaten erfasst.
- Datenerfassung durch

- **Beschleunigungssensor**

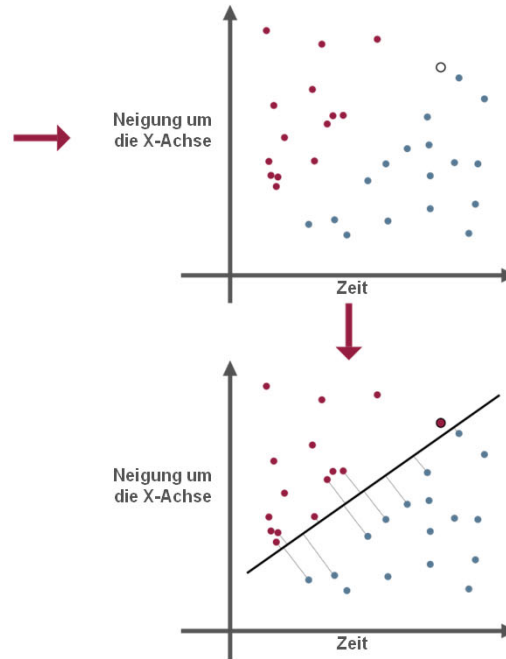
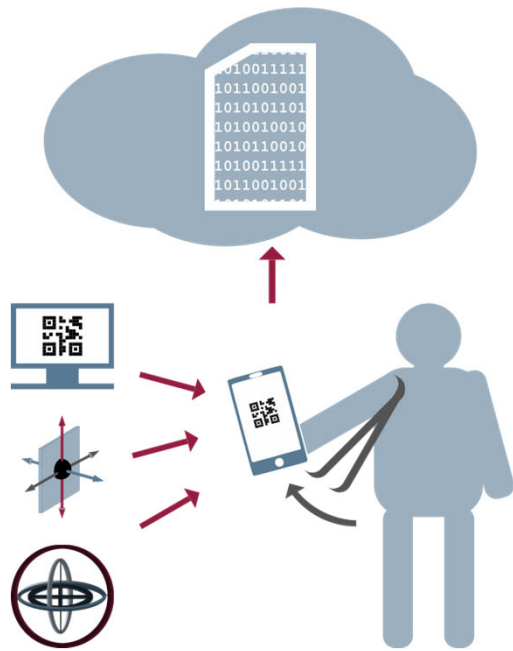


- **Lagesensor**



# Passive Authentifikation

## → Support-Vector-Machine (SVM)



Max Mustermann

### ■ Input-Daten:

- Nutzer holt Gerät aus Hosentasche
- Erfassen von **Lage** und **Beschleunigung** des Smartphones

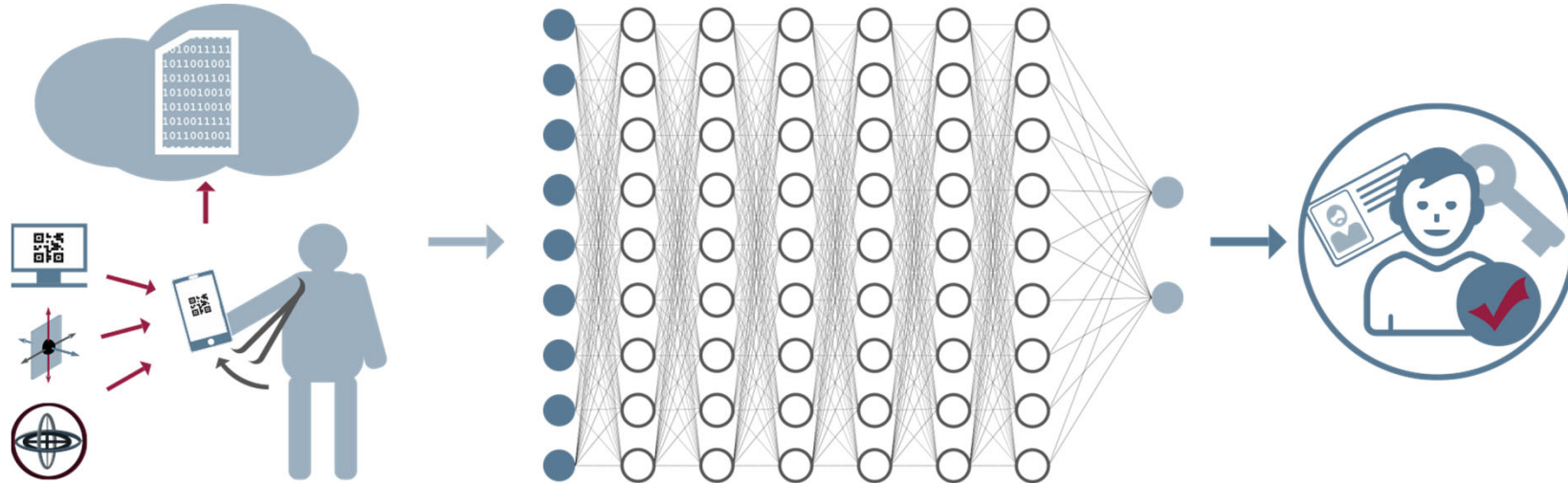
### ■ ML-Algorithmus:

- Daten werden anhand der Hyperebene/des Modell klassifiziert
- rote Übereinstimmung ist **positive** Klassifizierung
- blau eine **negative** Klassifizierung (bspw. anderer Nutzer)

### ■ Output:

- Authentisierung ist entweder erfolgreich oder schlägt fehl (**95 %**)

# Passive Authentifikation → Neuronales Netz



## Input-Daten:

- Lage und Beschleunigungsdaten des Nutzers werden erzeugt

## ML-Algorithmus:

- Eingabedaten werden in den künstlichen Neuronen in den Schichten verarbeitet

## Output:

Nutzer	Übereinstimmung
0	0,059 %
1	99,85 %
2	0,087 %

```
time, type, x, y, z
271, Accelerometer, -0.07606506, 9.173798, 3.6333618
277, Accelerometer, 1.0681152E-4, 9.146423, 3.5619507
279, Gyroscope, 0.027664185, 0.06774902, 0.02182006
...
```

```
[[5.9110398e-04 9.9853361e-01 8.7528664e-04]]
Predicted Class [1]
Predicted Person: Sandra Kreis
```

# Beispiele ethischer Herausforderungen

## → Privatheit vs. Allgemeinwohl

- Eine **Notwendigkeit** für die **Hinzunahme von personenbezogenen Daten** von Mitarbeitern könnte sich zum Beispiel in dem Fall ergeben, wenn ein **Energieversorger (Kritische Infrastruktur) angegriffen wird**.
- **Es besteht das Dilemma:**  
Die Ethik fordert, dass keine grundlegenden **Rechte wie die Privatheit** zugunsten eines höheren Ziels völlig aufgegeben werden dürfen.
- Dagegen steht die **Prämisse des Strebens nach dem Gesamtnutzen für die Gesellschaft**.
- Daraus resultiert die Fragestellung:  
„Wann ist es angeraten die **Rechte des Individuums** zugunsten des **Wohles für die Gemeinschaft** aufzuheben?“



# Beispiele ethischer Herausforderungen

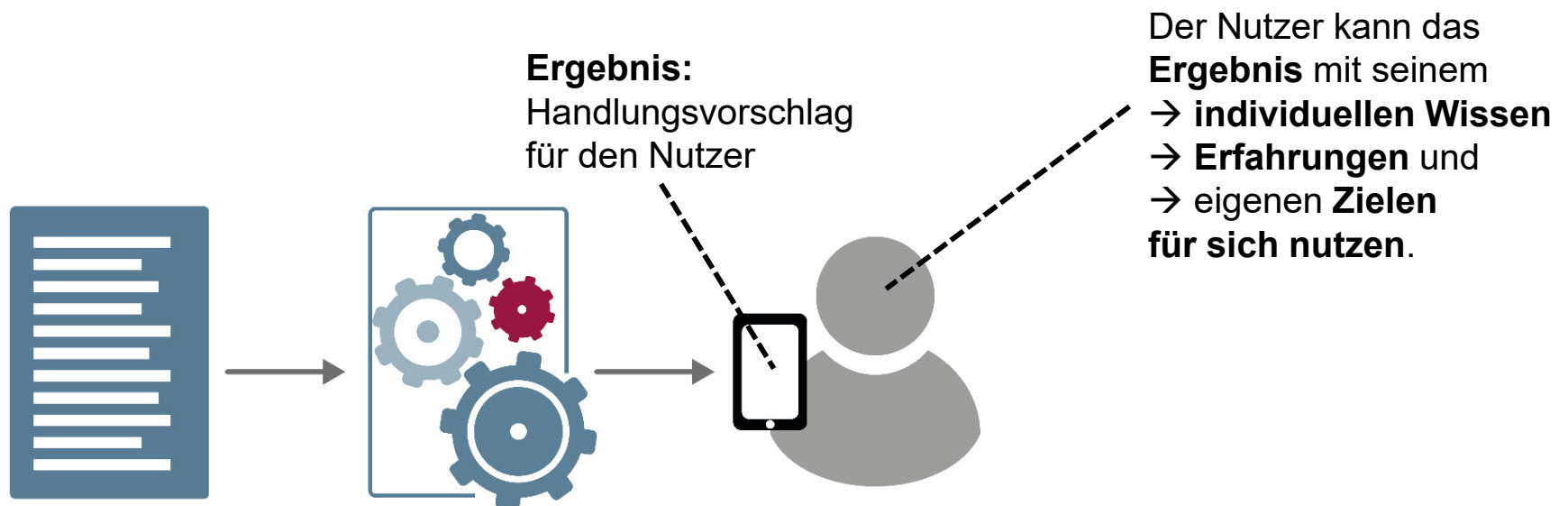
## → Neutralität vs. Voreingenommenheit (Bias)

- Die Auswahl von sicherheitsrelevanten Daten ist unter anderem durch Erfahrungswerte aus der Vergangenheit geprägt und beeinflusst die zukünftigen Ergebnisse.
- Das wird zum Beispiel deutlich an statistischen Analysen im Kontext der Polizeiarbeit in den USA.
- Es wird die Annahme zugrunde gelegt, dass eine bestimmte Bevölkerungsgruppe eher zu Straftaten neigt.
- Das legt den Fokus auf diese und wird dann kontinuierlich anhand aktualisierter Werte weiter bestätigt. Siehe „*Black Lives Matter*“
- Daraus ergibt sich die Fragestellung:  
„Wie kann die algorithmische Voreingenommenheit auf ein Minimum reduziert werden?“

# Beispiele ethischer Herausforderungen

## → Strike Back: Problem Unvollständigkeit der Daten

- Dem Konstrukt des ‚Strike Back‘ liegt die Hypothese zugrunde, dass **ein Angriff durch einen Gegenangriff beendet** werden kann.
- Die KI kann den Strike Back berechnen und automatisiert durchführen.
- Aber was ist, wenn dies gleichzeitig auch die Stromversorgung eines Krankenhauses lahmlegen würde?
- Die ethische Frage in diesem Kontext ist:  
**„Ist es möglich die Vollständigkeit der Daten zu erreichen und deren Relevanz zu beurteilen?“**



# Neue Strategien und Lösungen

→ Mehr **Zusammenarbeit** statt **Separation**

## Ungleichgewicht bei Angreifern und Verteidigern



## Zusammenarbeit hilft das Ungleichgewicht zu überwinden.

- **Echtzeitaustausch** von sicherheitsrelevanten Informationen (Daten) zwischen Wirtschaft, Staat, ... **motivieren** (*Vertrauen aufbauen, Kosten und Risiken reduzieren, ...*)
- **Erfahrungen** von Reaktionen (Daten) auf Angriffe **austauschen**
- **Gemeinsame Reaktionen** auf Angriffe **umsetzen**, um höhere Effekte zu erzielen und Schäden zu reduzieren ...
- **Gemeinsame** Definition und Umsetzung von notwendigen **Cyber-Sicherheitsmechanismen fördern**
- Notwendigkeit einer **Ethik-Diskussion** bei der Nutzung von KI, um eine hohe Akzeptanz zu erreichen.
- ...



**Westfälische  
Hochschule**

Gelsenkirchen Bocholt Recklinghausen  
University of Applied Sciences

## **Künstliche Intelligenz (KI) und Cyber-Sicherheit**

→ **Cyber-Sicherheit** braucht **Künstliche Intelligenz**

→ **Künstliche Intelligenz** braucht **Cyber-Sicherheit**

*Wir brauchen die **KI** in der **Cyber-Sicherheit**.*

*Aber: Wir müssen auch die **Risiken** der **KI** aktiv **reduzieren!***

Prof. Dr. (TU NN)

**Norbert Pohlmann**

*Professor für Informationssicherheit*

*Leiter des Instituts für Internet-Sicherheit - if(is)*

*Vorstandsmitglied des Verbands der Internetwirtschaft - eco*

*Vorstandsvorsitzender des Bundesverbands IT-Sicherheit - TeleTrusT*

**if(is)**  
internet-sicherheit.

## Wir empfehlen

- **Cyber-Sicherheit**

Das **Lehrbuch** für Konzepte, Mechanismen, Architekturen und Eigenschaften von Cyber-Sicherheitssystemen in der Digitalisierung“, Springer Vieweg Verlag, Wiesbaden 2019  
<https://norbert-pohlmann.com/cyber-sicherheit/>



- **7. Sinn im Internet (Cyberschutzraum)**

<https://www.youtube.com/cyberschutzraum>



- **Master Internet-Sicherheit**

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



## 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/_ifis)

<https://twitter.com/ProfPohlmann>

### YouTube

<https://www.youtube.com/user/InternetSicherheitDE/>

### Prof. Norbert Pohlmann

<https://norbert-pohlmann.com/>

## Quellen Bildmaterial

Eingebettete Piktogramme:

- Institut für Internet-Sicherheit – if(is)

## 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/>

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

D. Petersen, N. Pohlmann: „Kommunikationslage im Blick - Gefahr erkannt, Gefahr gebannt“, IT-Sicherheit – Management und Praxis, DATAKONTEXT-Fachverlag, 4/2014

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: „Diskriminierung und weniger Selbstbestimmung? Die Schattenseiten der Algorithmen“, tec4u, 12/17

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

N. Pohlmann: **Lehrbuch „Cyber-Sicherheit“**, Springer Vieweg Verlag, Wiesbaden 2019  
ISBN 978-3-658-25397-4s

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