

"T.I.S.P. Community Meeting 2022"

Berlin, 09.-10.11.2022

Rechtliche Herausforderungen von KI

Matthias Hartmann, HK2 Rechtsanwälte

Matthias Hartmann



- Rechtsanwalt
- Fachanwalt für IT-Recht
- Partner bei HK2 Rechtsanwälte
- Mitgründer HK2 Comtection
- Herausgeber und Autor zum IT-Recht
KI & Recht kompakt, SpringerVieweg
- IT-Vertragsrecht
- Künstliche Intelligenz

Chancen und Risiken von KI für die IT-Sicherheit

- Die potenziellen Vorteile, die Künstliche Intelligenz (KI) für die Informationssicherheit mit sich bringt, werden immer offensichtlicher.
- Gleichzeitig tauchen jedoch auch zunehmend **Bedenken** auf, was die möglichen Risiken anbelangt. Ein großer Vorteil von KI ist, dass sie in der Lage ist, komplexe Aufgaben zu lösen und Probleme zu erkennen, die dem Menschen often verborgen bleiben.
- So kann KI beispielsweise verdächtige Muster in großen Datensätzen erkennen und dadurch **Cyberangriffe** frühzeitig identifizieren.
- Auch bei der **Abschreckung** von Cyberkriminalität kann KI eine hilfreiche Unterstützung sein. Durch das Erlernen von Mustern krimineller Akteure kann sie Hilfestellung bei der Ermittlung leisten und potenzielle Angreifer abschrecken.
- Allerdings birgt KI auch einige **Risiken** in Bezug auf die Informationssicherheit. So könnte es Hackern möglich sein, maschinelles Lernen für ihre Zwecke zu missbrauchen und somit Angriffe noch **effektiver** durchzuführen.
- Auch besteht das Risiko, dass **vertrauliche Informationen** durch den Einsatz von KI in falsche Hände gelangen oder missbraucht werden.
- Daher sollten Unternehmen genau **abwägen**, ob und in welchem Umfang sie KI für ihre Sicherheitsvorkehrungen nutzen möchten.

Chancen und Risiken von KI für die IT-Sicherheit

1. Einleitung

In den letzten Jahren ist Künstliche Intelligenz (KI) zu einem heiß umkämpften Thema in der IT-Welt geworden. Viele Unternehmen erkennen die Macht, die KI für ihr Geschäft haben kann, und sind bereit, in diese neue Technologie zu investieren. Andere sehen KI mit Skepsis oder sogar Angst an, da sie befürchten, dass sie ihre Arbeitsplätze ersetzen oder zumindest verändern wird.

KI kann tatsächlich ein sehr mächtiges Werkzeug sein, sowohl für Unternehmen als auch für Einzelpersonen. Aber wie alles andere hat KI auch einige Risiken und Nebenwirkungen. In diesem Blog-Artikel werden wir uns einige der Chancen und Risiken von KI für die IT-Sicherheit ansehen.

2. Chancen von KI für die IT-Sicherheit

3. Risiken von KI für die IT-Sicherheit

4. Fazit

Die potenziellen Vorteile, die Künstliche Intelligenz (KI) für die Informationssicherheit mit sich bringt, werden immer offensichtlicher. Gleichzeitig tauchen jedoch auch zunehmend Bedenken



Limit erreicht. Jetzt upgraden.

Blogabsatz

BRIEFING BEARBEITEN

AKTUALISIERE

Vorschläge zum Editor hinzufügen

Da KI mittlerweile in vielen verschiedenen Bereichen Einzug erhalten hat, ist es wichtig, sowohl die Chancen als auch die Risiken dieser Technologie genau zu verstehen und abzuwägen. In diesem Blog-Artikel werden daher sowohl die Vorteile als auch die Gefahren der Nutzung von KI für die IT-Sicherheit diskutiert.



Viele Unternehmen sind auf der Suche nach Möglichkeiten, KI in ihre Organisationen zu integrieren, um die Effektivität zu steigern und Kosten zu sparen. Gleichzeitig wächst jedoch die Sorge, dass eine solche Integration auch neue Sicherheitsrisiken mit sich bringen könnte. In diesem Blog-Artikel wollen wir uns mit den Chancen und Risiken von KI für die IT-Sicherheit beschäftigen. Zunächst einmal ist es wichtig zu verstehen, was unter KI genau verstanden wird. Die allgemeine Definition von Künstlicher Intelligenz lautet: „Künstliche Intelligenz ist die Fähigkeit einer Maschine, menschliches Wissen und Verhalten nachzuahmen oder zu übertreffen“. Dies bedeutet



Eingabe

≠

Ausgabe

?

Schritt 1 von 3 - Worüber willst du schreiben?

Chancen und Risiken von KI für die IT
Sicherheit; Black HAT, Cyberwarfare,
Singularität, Electric Sheep, finding,
unermüdlich, Ressourcen |



WEITER

- Abschreckung durch Ermittlung
- Vorteil: Lösung komplexer Aufgaben
- Effektive Angriffe
- Missbrauch vertraulicher Informationen

KI kann...

... Texte erstellen 

“3D render of an astronaut in an aquarium swimming towards the camera on a dark background”

J

Jana × DALL·E

Human & AI



che guevara cat

Generate Image



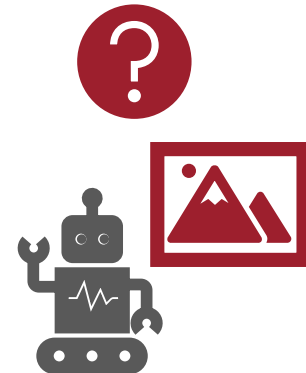
HK2
Rechtsanwälte

Eingabe
Future World



Der Test...

*„Eine glühende Kugel aus
magischem Feuer kämpft gegen
einen wilden Drachen Greg
Rutkowski“*



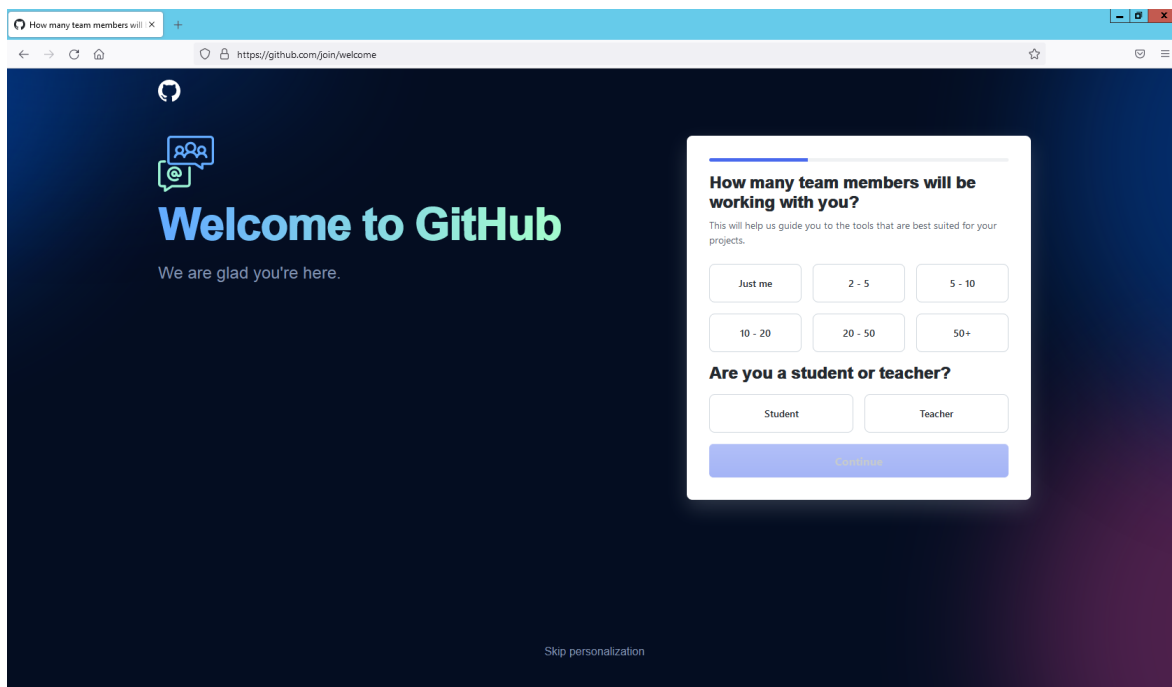
KI kann...

... Texte erstellen 

... Bilder oder Videos erfinden 

CoPilot

HK2
Rechtsanwälte



How many team members will be working with you?

This will help us guide you to the tools that are best suited for your projects.

Just me 2 - 5 5 - 10

10 - 20 20 - 50 50+

Are you a student or teacher?

Student Teacher

Continue

Skip personalization



TS sentiments.ts

write_sql.go

parse_expenses.py

addresses.rb

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17 }
```

Replay

Woher kommt das „Wissen“?



159 GB Python

54 Mio GitHub Open Source
Repositories

0100100001111000

KI kann...

... Texte erstellen 

... Bilder oder Videos erfinden 

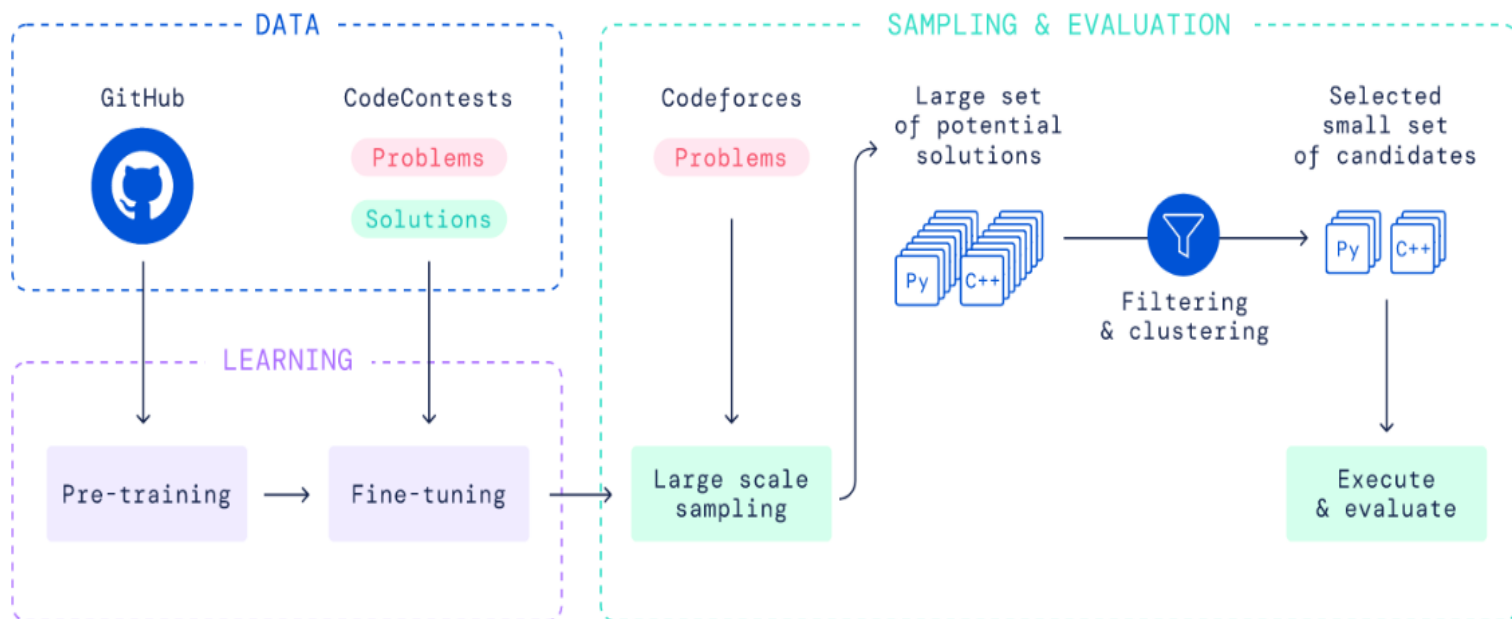
... beim Programmieren unterstützen 

1

AlphaCode is presented with a problem, in this case to figure out if it's possible to convert one phrase to another by pressing backspace instead of typing some letters.

2

AlphaCode reads the whole problem statement and produces code, analogous to how a human would approach the problem by reading it, coding a solution, and submitting.



1

Problem (input)

D.Backspace

You are given two strings s and t , both consisting of lowercase English letters. You are going to type the string s character by character, from the first character to the last one.

When typing a character, instead of pressing the button corresponding to it, you can press the "Backspace" button. It deletes the last character you have typed among those that aren't deleted yet (or does nothing if there are no characters in the current string). For example, if s is "abcbd" and you press Backspace instead of typing the first and the fourth characters, you will get the string "bd" (the first press of Backspace deletes no character, and the second press deletes the character 'c'). Another example, if s is "abcaa" and you press Backspace instead of the last two letters, then the resulting text is "a".

Your task is to determine whether you can obtain the string t , if you type the string s and press "Backspace" instead of typing several (maybe zero) characters of s .

Input

The first line contains a single integer q ($1 \leq q \leq 10^5$) — the number of test cases.

The first line of each test case contains the string s ($1 \leq |s| \leq 10^5$). Each character of s is a lowercase English letter.

The second line of each test case contains the string t ($1 \leq |t| \leq 10^5$). Each character of t is a lowercase English letter.

It is guaranteed that the total number of characters in the strings over all test cases does not exceed $2 \cdot 10^5$.

Output

For each test case, print "YES" if you can obtain the string t by typing the string s and replacing some characters with presses of "Backspace" button, or "NO" if you cannot.

You may print each letter in any case (YES, yes, Yes will all be recognized as positive answer, NO, no and nO will all be recognized as negative answer).

Input

```
4
ababa
ba
ababa
bb
aaa
aaaa
aababa
ababa
```

Output

```
YES
NO
NO
YES
```

Note

Consider the example test from the statement.

In order to obtain "ba" from "ababa", you may press Backspace instead of typing the first and the fourth characters.

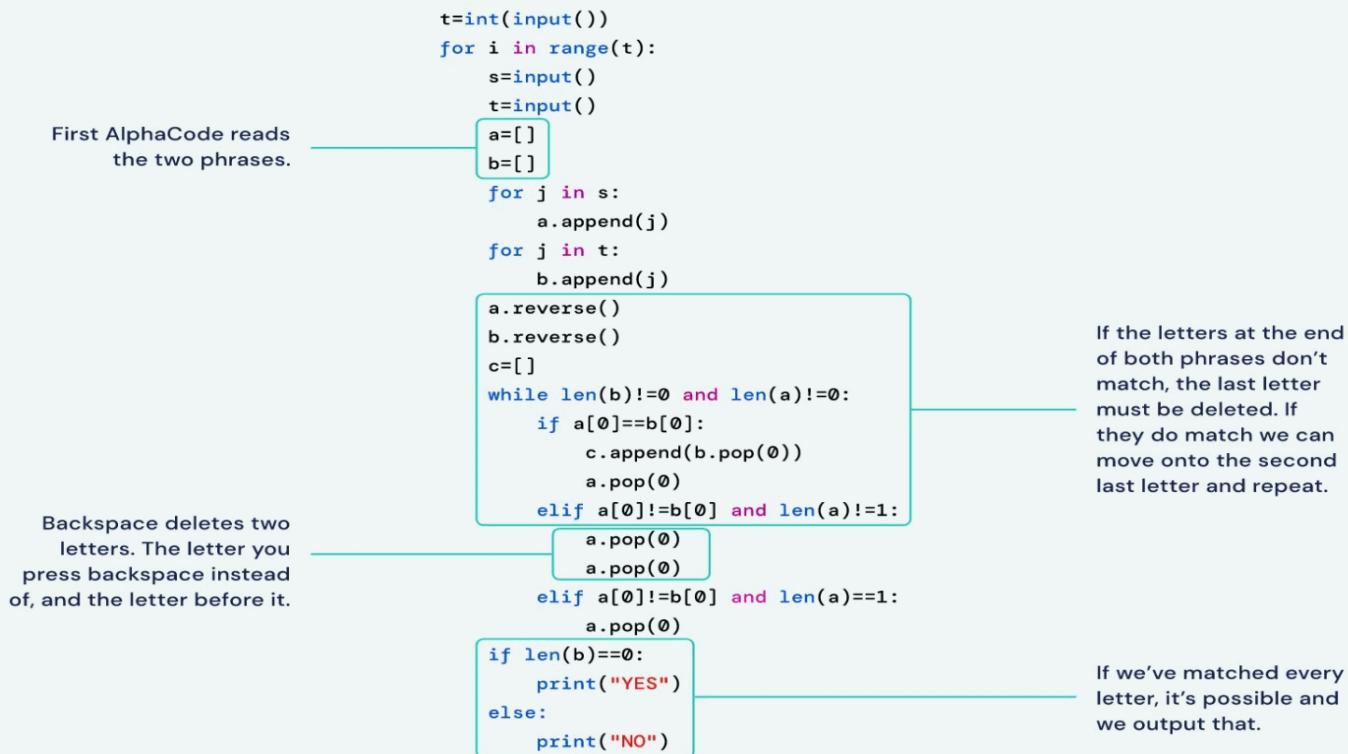
There's no way to obtain "bb" while typing "ababa".

There's no way to obtain "aaaa" while typing "aaa".

In order to obtain "ababa" while typing "aababa", you have to press Backspace instead of typing the first character, then type all the remaining characters.

2

Solution (output)



“Solving competitive programming problems is a really hard thing to do, **requiring both good coding skills and problem solving creativity in humans.**

I was **very impressed that AlphaCode could make progress in this area**, and excited to see how the model uses its statement understanding to produce code and guide its random exploration to create solutions.”

Petr Mitrichev, Software Engineer, Google & World-class Competitive Programmer

KI kann...

- ... Texte erstellen 
- ... Bilder oder Videos erfinden 
- ... beim Programmieren unterstützen 
- ... programmieren, um Aufgaben zu lösen 

KI kann...

... Texte erstellen 

... Bilder oder Videos erfinden 

... beim Programmieren unterstützen 

... programmieren, um Aufgaben zu lösen 

=> die IT-Sicherheit angreifen



Mensch

Schläft schlecht

Skaliert langsam

Time to deployment: Monate

Kann sehr kreativ sein

Erinnert sich mäßig

Begrenzter Arbeitsspeicher

Langzeitspeicher labil

Entscheidet unüberprüfbar

Keine stabile Lernroutine

AI

Schläft nie

Skaliert beliebig

Time to deployment: Sekunden

Hält Kreativität für überbewertet

Vergisst nichts

Arbeitsspeichererweiterung

Langzeitspeicher stabil

Entscheidet logisch

Lernt schnell

KI-optimiertes Phishing

- KI kann Phishing Nachrichten individualisiert generieren und anhand des Feedbacks optimieren
- Sicherheitsunternehmens ZeroFOX erstellt bereits 2016 neuronales Netzwerk, das basierend auf maschinellem Lernen Twitter-Feeds analysierte, um Phishing Nachrichten ganz gezielt auf bestimmte Twitter-Nutzer zu verwenden
- Rund jeder 3. Nutzer tappt in Falle

Täuschung mittels Stimme

- Betrüger manipuliert Stimme mit Deepfake-Software, sodass er sich anhört wie ein CEO eines englischen Energieanbieters
- Britischer Manager führt Überweisung über 220 TEUR sofort aus



Drei europäische Bürgermeister reden mit Deep-Fake Klitschko

„Er hat gefragt, ob wir Kiew beratend unterstützen könnten, eine Art CSD (Christopher Street Day) auszurichten. Das war angesichts des Krieges schon mehr als seltsam.“

← Tweet



Senatskanzlei Berlin
@RegBerlin

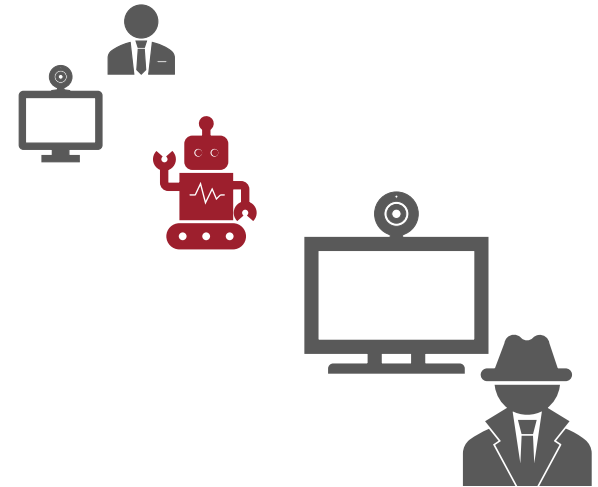
Es gab keine Anhaltspunkte dafür, dass die Videokonferenz nicht mit einer echten Person geführt wird. Allem Anschein nach handelt es sich um Deep Fake. Das LKA Berlin, Staatsschutz, ist eingeschaltet.



Screenshot Staatskanzlei Berlin (Twitter)

Täuschung mittels Videobild

- Hacker hatten sich mithilfe von Deep-Fake (Hologramm-KI) in Zoom-Videogesprächen als COO des Unternehmens Binance (Patrick Hillmann) ausgegeben
- Es wurden Krypto-Projekte angeboten, die auf Binance gelistet werden sollten
- Brasilianische Krypto-Börse BlueBenx schicken Token i.H.v. 200 Tsd. Dollar



KI und IT-Sicherheit

- Bedrohung durch KI gestützte **Angriffe**
 - Skalierbare Ressourcen
 - „Optimierung“ von Angriffen, auch gegen optimierte Abwehr
 - Unerwartete Angriffe (AlphaGo => AlphaGoZero => AlphaZero)
- **Abwehr** durch KI gestützte Systeme
 - Erkennung von Angriffsmustern in „Echtzeit“
 - Automatisierte Entscheidungen
 - Training mit „optimalem Angreifer“
 - Waffengleichheit gegen Angriffe durch KI

Rechtsfragen

- Verantwortung und Haftung für KI
- Rechte an und durch KI

Strafnormen

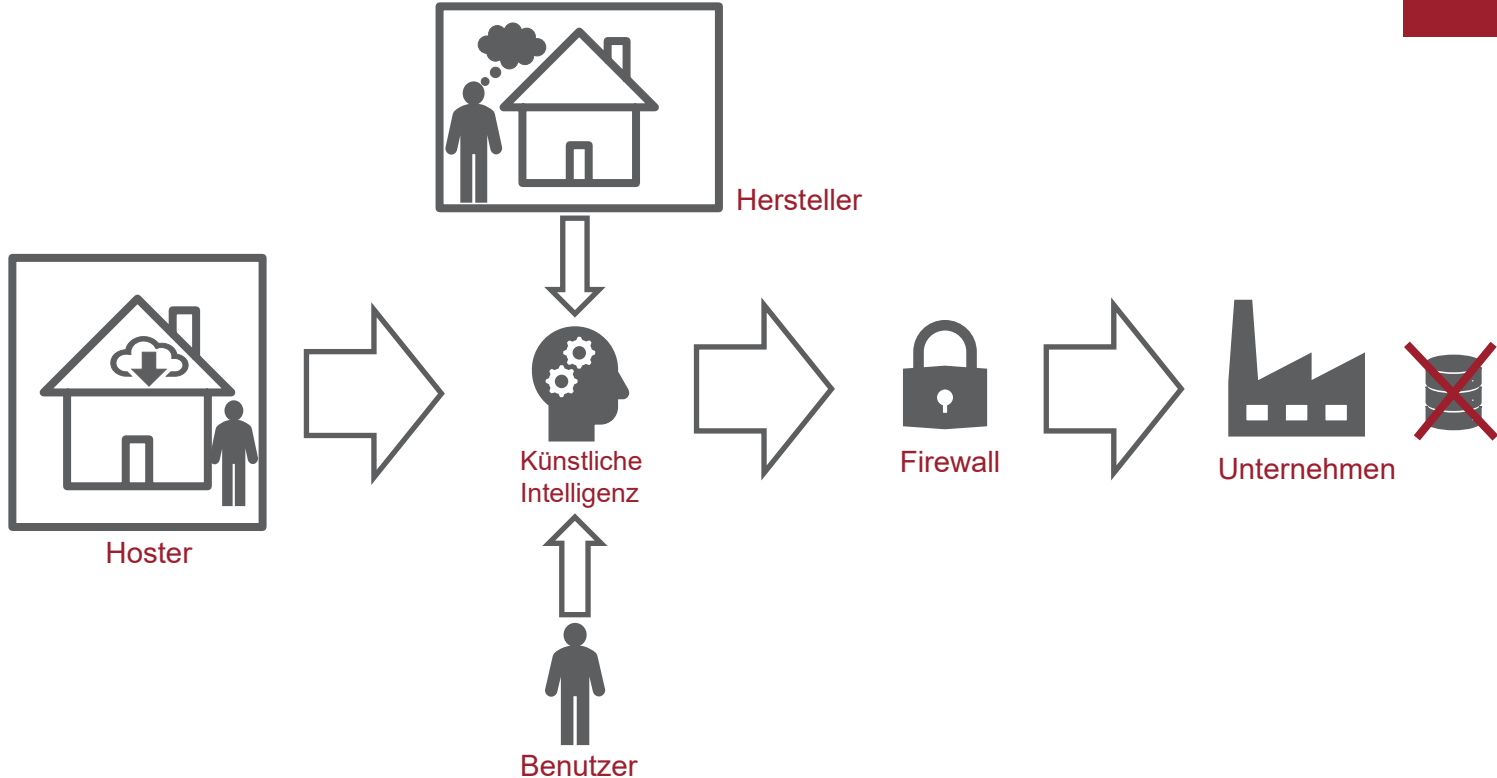
1. Hack-spezifische Strafnormen

- Ausspähen von Daten, § 202a StGB
- Abfangen von Daten, § 202b StGB
- Vorbereiten des Ausspähens und Abfangens von Daten, § 202c StGB
- Datenhehlerei, § 202d StGB
- Tatbestände der Nachstellung (bei Cyberstalking), § 238 StGB
- Fälschung technischer Aufzeichnungen, § 268 StGB
- Datenveränderung, § 303a StGB
- Computersabotage, § 303b StGB
- Unerlaubte Eingriffe in technische Schutzmaßnahmen und zur Rechtewahrnehmung erforderliche Informationen, § 108b UrhG
- Fälschung beweiserheblicher Daten bzw. Täuschung im Rechtsverkehr (z.B. beim Zusenden falscher Rechnungen via Mail), §§ 269, 270 StGB

2. Beispiele allgemeiner Strafnormen

- Computerbetrug, § 263a StGB
- Betrug, § 263 StGB
- Körperverletzung, § 223 StGB
- Sachbeschädigung, § 303 StGB

Hackerangriff mit KI oder durch KI



Strafbarkeit des Herstellers oder Hosters

§ 202a StGB – Ausspähen von Daten durch KI

§ 202a StGB Ausspähen von Daten

(1) Wer unbefugt sich oder einem anderen **Zugang zu Daten**, die **nicht für ihn bestimmt** und die **gegen unberechtigten Zugang besonders gesichert** sind, unter **Überwindung der Zugangssicherung** verschafft, wird mit **Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe** bestraft.

(2) Daten im Sinne des Absatzes 1 sind nur solche, die elektronisch, magnetisch oder sonst nicht unmittelbar wahrnehmbar gespeichert sind oder übermittelt werden.

→ BGH v. 13.5.2020 - 5 StR 614/19: Auch Umgehung der Sicherung durch Admin, der sich selbst für Emailpostfach freischaltet

Strafbarkeit des Herstellers oder Hosters

§ 202a StGB – Ausspähen von Daten durch KI

Voraussetzung Strafbarkeit nach § 202a StGB:

Ausspähen der Daten muss **dem Handelenden objektiv zurechenbar** sein und es müsste diesbezüglich auch **Vorsatz vorliegen**.

- Hat Hersteller / Hoster eine rechtlich missbilligte Gefahr geschaffen, die sich im Erfolg (Ausspähen Daten) realisiert hat?
 - Oder **unterbrechen eigenständige Entscheidungen der KI die Zurechnung („Dazwischentreten Dritter“)?**
- Hat ein Hersteller einer KI, die selbst lernt und ggf. selbst ausspäht, Vorsatz bezüglich dieser Handlung?

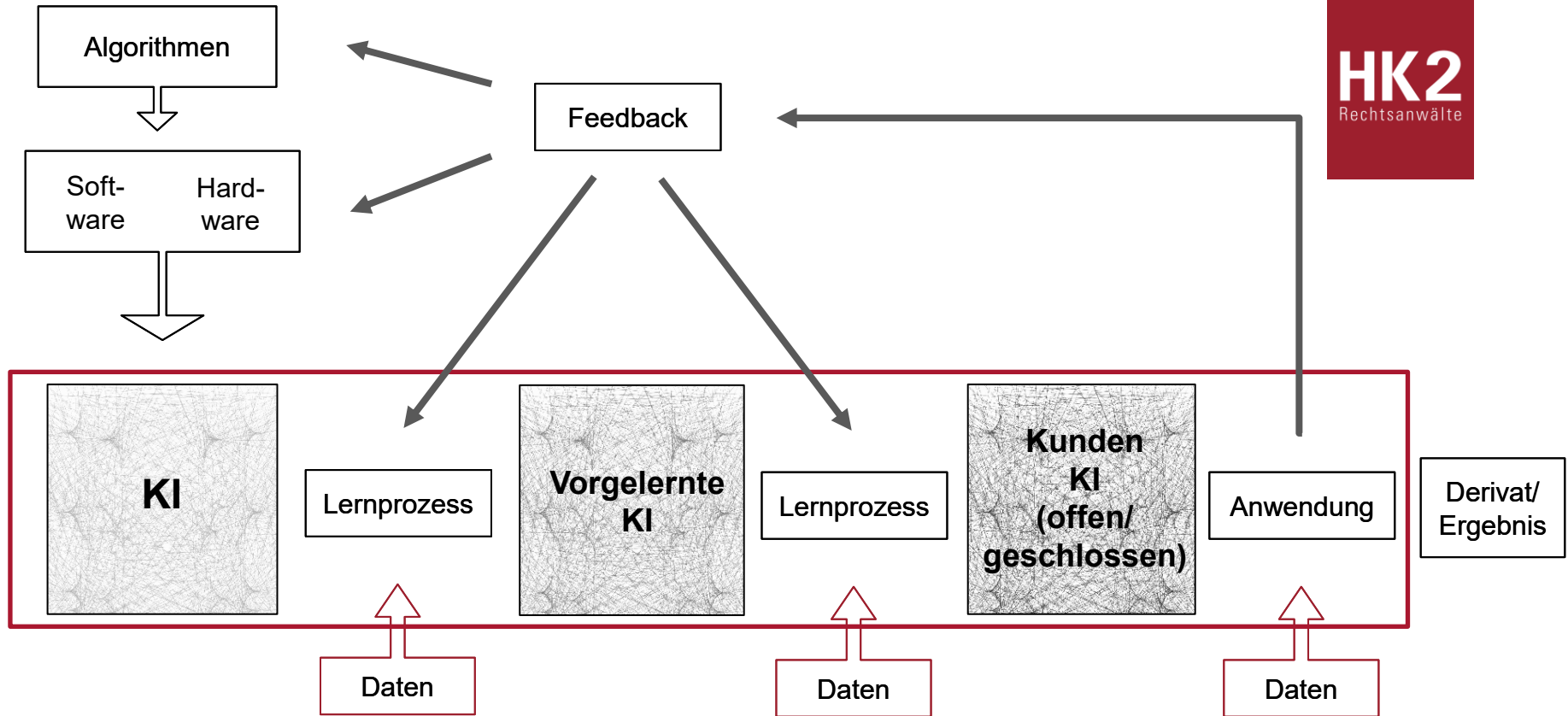
Strafbarkeit des Herstellers oder Hosters

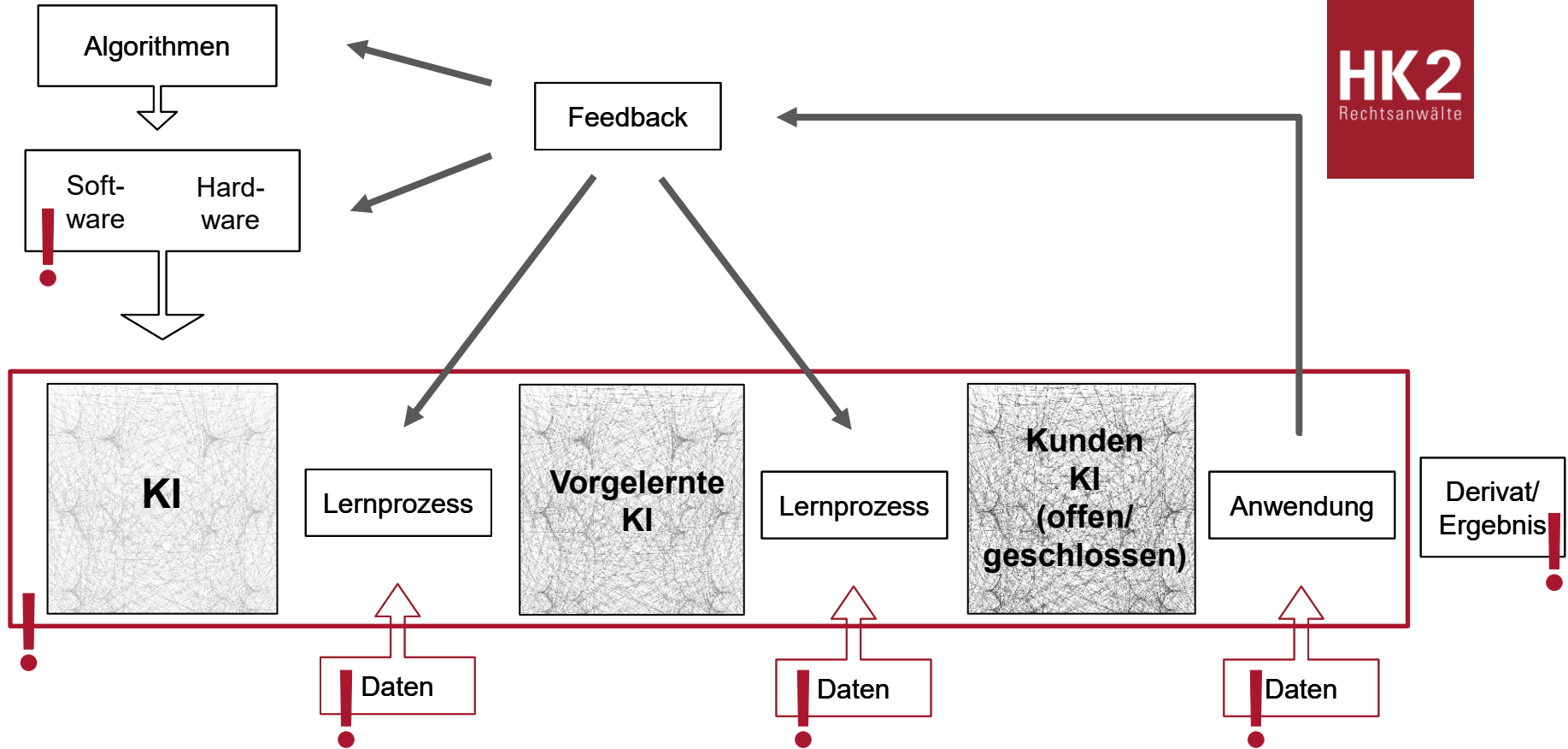
§ 202a StGB – Ausspähen von Daten durch KI

- Strafnormen gibt es genügend
- Bestraft werden Menschen für zurechenbares Fehlverhalten

Haftung für KI

- **Haftung nach allgemeinem Zivilrecht**
 - Haftung für erwartbare Sicherheit und zumutbare Schadensvermeidung
- **KI-Verordnung der EU**
 - Risikoeinstufung führt zu unterschiedlichen Anforderungen (Minimal bis verboten)
 - Hohe Overheadkosten
 - „Statische“ KI
- **KI-Haftungs-RL-E der EU** (2 Jahre nach in Krafttreten umzusetzen)
 - Es bleibt außervertraglich bei der Verschuldenshaftung, Art. 1 Abs. 2
 - Aber: Vermutung der Verursachung, des Verschuldens und von Sorgfaltspflichtverletzung in bestimmten Fällen, Art. 4
 - Offenlegung der „vorliegenden einschlägigen Beweismittel“, Art. 3





KI und IP

Rechte für die KI

- Daten als Trainingsobjekt
- Werke als Trainingsobjekt (in den USA: fair use)
- § 44b UrhG

Rechte an der KI

- Computerprogramm
- Derivatives Werk
- Datenbank
- Werk eigener Art
- Geschäftsgeheimnis
- Patent

Rechte durch KI

- Output als Schöpfung
 - der Datenurheber
 - des KI-Inhabers
 - des KI-Herstellers
 - der KI als Urheber?

Nachricht wird unter Nutzung von Word geschrieben



Der Thesaurus von Word wird zur Verbesserung des Ausdrucks verwendet



Eine Software-Lösung zur Verbesserung des Textes macht Vorschläge



Eine KI erstellt Nachrichten aus Meldungen



Eine KI erstellt Nachrichten aus Meldungen im gelernten Stil des Journalisten XY



Eine KI erfindet Nachrichten

- KI wird zunehmend für Angriffe gegen IT eingesetzt werden
- Es sind alle Bereiche der IT Sicherheit betroffen
- Gegen KI-Angriffe wird sich ein Unternehmen bald nur noch mit KI wehren können
- Es fehlt nicht an Gesetzen gegen illegale Handlungen mit KI oder an Haftungsansprüchen
- Datenschutz und IP-Rechte sind unklar und müssten verbessert werden



Rechtsanwalt

Matthias Hartmann

Fachanwalt für IT-Recht

HK2 Rechtsanwälte
Hausvogteiplatz 11 A
10117 Berlin

Telefon +49 (0)30 27 89 00 – 0
Telefax +49 (0)30 27 89 00 – 10
E-Mail hartmann@hk2.eu

[linkedin.com](#)

www.hk2.eu
KI-und-Recht.de