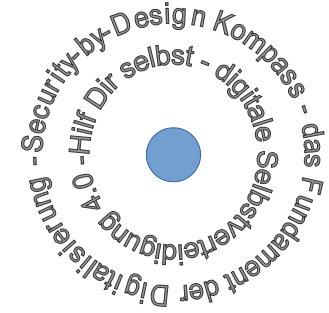


5. Erklärbarkeit [1]

- **Nachvollziehbarkeit**
 - **Erklärbarkeit**
 - **Interpretierbarkeit**
 - (→) **Transparenz**
- Nicht nur die Erklärbarkeit des Ergebnisses ist hier gefordert, auch und vor allem die Nachvollziehbarkeit und Erklärbarkeit des Prozesses der Entscheidungsfindung.

- **transparente** Prozesses
- Fähigkeiten und Zweck von KI offen kommunizieren
- Entscheidungen in größtmöglichem Umfang erklärbar machen
- keine „BlackBox“-Algorithmen



Die KI „AlphaGo“ besiegte im März 2016 den weltbesten Profispieler des Brettspiels Go in drei Spielen hintereinander.

<https://www.heise.de/newsticker/meldung/Kuenstliche-Intelligenz-AlphaGo-Zero-uebertrumpft-AlphaGo-ohne-menschliches-Vorwissen-3865120.html>

AlphaGo basiert auf maschinellem Lernen und der Verwendung neuronaler Netze. Letztere sind von der Funktion des menschlichen Gehirns inspiriert und sollen dieses imitieren. Entsprechend können sie beliebig komplex werden, was zur Folge hat, dass die Nachvollziehbarkeit bzw. Erklärbarkeit zunehmend erschwert wenn nicht gar unmöglich wird.

Eine andere Frage zur **Transparenz**: Ist meine Kontaktperson bzw. mein Gegenüber ein Mensch oder eine KI? (Problem: Verdeckte KI-Systeme) [3]

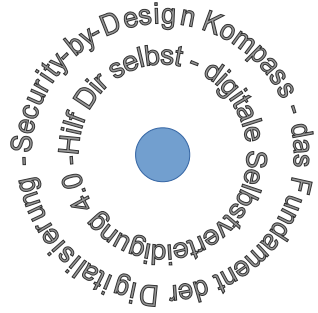
Hambacher Erklärung [2]

I. Künstliche Intelligenz und Datenschutz

Personenbezogene Daten in den falschen Händen können missbraucht werden.

Verarbeitung personenbezogener Daten durch KI birgt Risiken für Rechte und Freiheiten von Menschen. Bsp.: KI-Bewertung von Bewerbungsunterlagen. Trainiert wurde das System ausschließlich mit Daten männlicher Bewerber. → Weibliche Bewerber im echten Betrieb wurden sehr viel schlechter eingeschätzt, da sie dem System gänzlich unbekannt waren.

→ politische, gesellschaftliche und rechtliche Begleitung der KI



III. Entwicklung von KI bedarf Steuerung

Überwachung der Anwendung des Datenschutzrechts.

Hohe Dynamik in der Entwicklung von KI → Grenzen dieser Entwicklung noch gar nicht absehbar.

→ Risiken nicht pauschal einschätzbar. Begleitung der Entwicklung und des Einsatzes künstlicher Intelligenz durch Politik, Anwender, Wissenschaft und Datenschutzaufsichtsbehörden sowie die Beachtung **ethischer Grundsätze**.

Hambacher Erklärung [2]

II. Datenschutzrechtliche Anforderungen an künstliche Intelligenz

1. Mensch nicht als Objekt

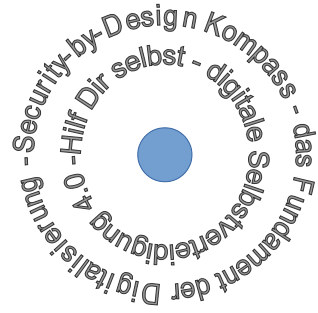
Garantie der Würde des Menschen insbesondere im Fall staatlichen Handelns. Vollständig automatisierte Entscheidungen oder Profiling durch KI nur eingeschränkt zulässig.

2. Nur für verfassungsrechtlich legitimierte Zwecke

Auch Zweckbindung darf nicht aufgehoben werden.

3. Transparent, nachvollziehbar, erklärbar (Siehe (5))

4. **Vermeidung von Diskriminierungen**
Hohe Abhängigkeit von Trainingsdaten. Abgebildete Diskriminierungen in Trainingsdaten übertragen und verfestigen sich beim Einsatz von KI. Bewertung von Risiken für Rechte und Freiheiten von Personen.



5. Datenminimierung

Große Bestände von Trainingsdaten → für personenbezogene Daten gilt: Datenminimierung. Nicht mehr Daten verarbeiten als zwangsläufig notwendig. Möglicherweise Verarbeitung vollständig anonymisierter Daten für den Zweck bereits ausreichend?

6. Verantwortlichkeit

Ermittlung und Kommunikation von Verantwortlichkeiten. Gewährleistung rechtmäßiger Verarbeitung, Betroffenenrechte, Sicherheit der Verarbeitung und der Beherrschbarkeit des Systems. Auch Verhindern von Manipulationen durch Dritte.

7. Technische und Organisatorische Standards

Bisher keine speziellen Standards und Maßnahmen für datenschutzkonformen Einsatz von KI. Diese sind jedoch für die Sicherstellung einer datenschutzgerechte Verarbeitung zutreffen.

Was sagt die EU? Wichtige Richtlinien: [3]

(1) Menschliche Autonomie, Schadensverhütung, Fairness und Erklärbarkeit

(2) Berücksichtigung besonders schutzbedürftiger Gruppen wie Kinder oder Menschen mit Behinderungen

(3) Bewusstsein für Risiken beim Einsatz von KI

(4) Vorrang menschlichen Handelns, menschliche Aufsicht, technische Robustheit und Sicherheit, Privatsphäre und Datenqualitätsmanagement, Transparenz, Vielfalt, Nichtdiskriminierung und Fairness, gesellschaftliches und ökologisches Vorgehen sowie Rechenschaftspflicht.

(5) technische und nichttechnische Methoden zur Umsetzung der Anforderungen

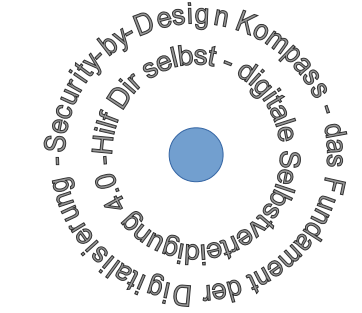
(6) Bewertung von KI-Systemen, Veröffentlichung von Ergebnissen, öffentlicher Diskurs

(7) Proaktive Informationsvermittlung an betroffene Kreise (Fähigkeiten und Grenzen von KI)

(8) Rückverfolgbarkeit und Nachprüfbarkeit von KI-Systemen, insb. in kritischen Zusammenhängen und Situationen

(9) Beteiligung der Interessenträger am Lebenszyklus des Systems Schulungen, Ausbildungsförderung, Kompetenzen

(10) Kompromisse zwischen den Anforderungen im Fall von wahrscheinlich auftretenden Spannungen dazwischen.



(11) Bewertungsliste für vertrauenswürdige KI in Bezug auf Entwicklung, Einführung und Nutzung von KI-Systemen

(12) Bewertungsliste kann niemals erschöpfend sein. Kein Abhaken von Punkten einer Liste, sondern kontinuierlicher Prozess von Ermittlung und Umsetzung von Anforderungen, Bewertung von Lösungen und Erzielung besserer Ergebnisse.

[3]