

Ein Rechtsrahmen für den Einsatz von Künstlicher Intelligenz (KI) in der Schweiz

KIRT – 18. Oktober 2024

Prof. Dr. Nadja Braun Binder | Prof. Dr. Florent Thouvenin

Herausforderungen

- **Transparenz**
- **Diskriminierung**
- **Datenschutz**
- **Manipulation**
- **Haftung**



Aktivitäten und Publikationen



Digital Society Initiative

Positionspapier

Ein Rechtsrahmen für Künstliche Intelligenz

Die grossen technischen Fortschritte im Bereich der **Künstlichen Intelligenz (KI)** und der Einsatz dieser Technologien in einer Vielzahl von Bereichen werfen grundlegende Fragen zu den Auswirkungen auf Individuen und die Gesellschaft auf. Der Begriff der Künstlichen Intelligenz weckt bisweilen irreführende Assoziationen und diffuse Ängste. Aus technischer Perspektive handelt es sich um einen etablierten Sammelbegriff, der **eine Reihe von Technologien** umfasst, die automatisierte Entscheidungen fällen, Empfehlungen machen, Schlussfolgerungen ziehen oder Vorhersagen treffen. Dazu gehören wissensbasierte Systeme und statistische Methoden ebenso wie Ansätze des maschinellen Lernens (z.B. unter Einsatz neuronaler Netze). Die grosse Leistungsfähigkeit dieser Technologien basiert meist auf der Aneinanderreihung einer Vielzahl von mathematischen Optimierungen, die unter Nutzung grosser Rechnerkapazitäten Strukturen aus grossen Datenmengen extrahieren.

Florent Thouvenin, Markus Christen, Abraham Bernstein, Nadja Braun Binder, Thomas Burri, Karsten Donnay, Lena Jäger, Mariela Jaffé, Michael Krauthammer, Melinda Lohmann, Anna Mätzener, Sophie Mützel, Liliane Obrecht, Nicole Ritter, Matthias Spielkamp, Stephanie Volz

Dieses Positionspapier wurde im Rahmen eines Workshops erarbeitet, der vom 26.–28. August 2021 in Balsthal durchgeführt und vom Strategy Lab der Digital Society Initiative (DSI) der Universität Zürich finanziert wurde. Neben den Autor*innen dieses Papiers haben auch drei Vertreter*innen der Bundesverwaltung an diesem Workshop teilgenommen, nämlich Monique Cossali Sauvain (BJ), Roger Dubach (EDA) und Thomas Schneider (BAKOM). Sie vertreten die Schweiz im Ad Hoc Komitee des Europarates zu Künstlicher Intelligenz (CAHAI). Weitere Informationen: dsi.uzh.ch/strategy-lab

➤ Bisherige Entwicklungen

- Bericht interdepartementale Arbeitsgruppe (2019)
- Positionspapier DSI (2021)
- White Papers ITSL / ePIAF (2024)
- Auslegeordnung Bundesrat (2024)

➤ Hauptpunkte

- Technologieneutrale Regelung
- Fokus auf konkrete Nachteile
- Kombination von allgemein anwendbaren («horizontalen») und sektorspezifischen («vertikalen») Regeln

Autonomer Nachvollzug AI Act?

➤ Argumente pro..

- Geringer Aufwand für die Schweiz: «copy paste», aber..
- Anerkennung von Konformitätsbewertungen, aber...

➤ .. und contra

- Keine technologieneutrale Regelung – nicht «future proof»
- AI Act = Produktsicherheitsrecht
 - Doppelungen mit sektorspezifischen Regelungen für Produktsicherheit
 - Konkretisierung in technischen Normen, die de facto auch in Schweiz relevant sein werden
- Regulierung stärkt die Marktposition von «Big Tech»

➤ Gründe der EU für AI Act

- Beschränkte Kompetenzen der EU: Umfassende Kompetenz für Binnenmarkt
- «Brussels Effect» und Ziel einen globalen Standard zu setzen

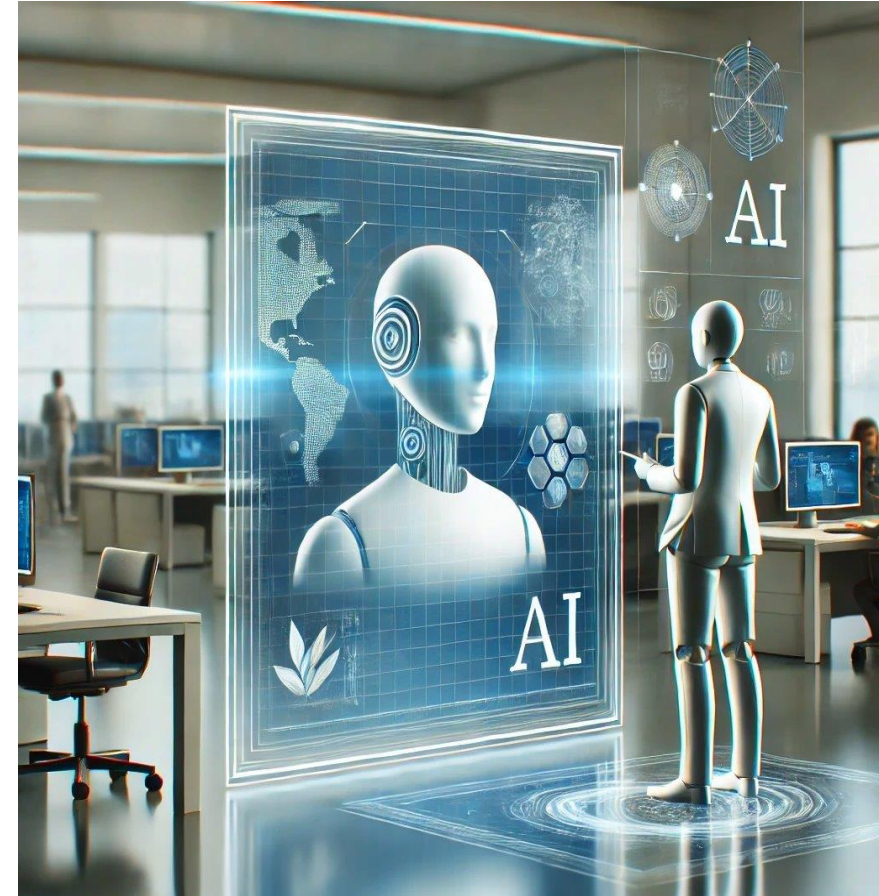


Transparenz durch öffentliches Verzeichnis

➤ Problemstellung

- Erhöhte Transparenz für KI-Einsatz durch Behörden
 - Ebene Bund: Motion 21.4508 Schlatter; Postulat 23.3566 Schlatter
 - Ebene Kantone: z. B. Kanton ZH Postulat 9/2022, § 13 Abs. 3 revidiertes IDG-ZH
- Rechtsstaatskonforme und diskriminierungsfreie digitale Transformation der Verwaltung
- Stärkung der Vertrauensbeziehung zwischen Staat und Bevölkerung
- Sichtbarmachen allfälliger Diskriminierungsprobleme

➤ Fokus «algorithmisches System (ADMS)»



Transparenz durch öffentliches Verzeichnis

➤ Ziele eines ADMS-Verzeichnisses

- Schutz des Menschen und der Menschenwürde
- Kontrolle staatlichen Handelns
- Vertrauensbildung
- Dokumentation

➤ Erfasste Systeme

- Abstellen auf die Auswirkungen der mithilfe des Systems getroffenen Entscheidung auf Rechte und Pflichten von Menschen
- Nicht erfasst: Systeme zur Erfüllung von administrativen Hilfstätigkeiten

➤ Regelungsort?

- Gesetz im formellen Sinn
- Zielführend: Verankerung im Datenschutzgesetz (DSG, SR 235.1)

Transparenz durch öffentliches Verzeichnis

➤ Regelungsgegenstand

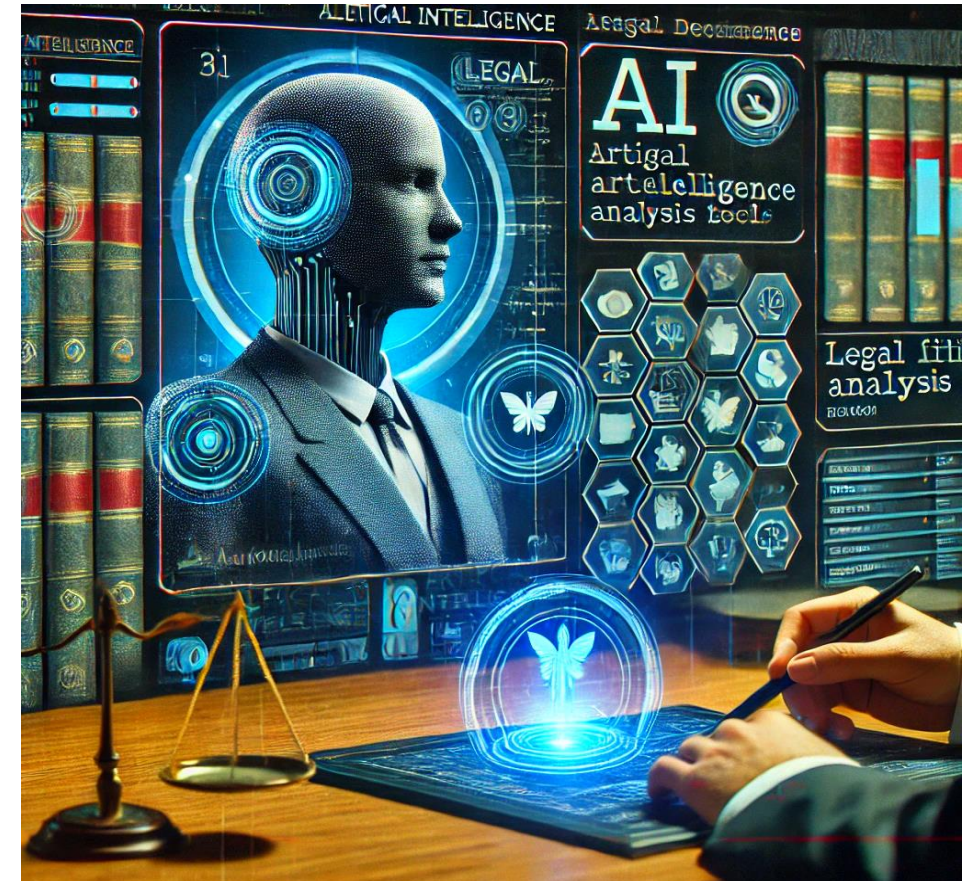
- Grundlegende Informationen im Gesetz im formellen Sinn:
 - die Stelle, die das Verzeichnis führt (z. B. EDÖB).
 - die zu erfassenden Systeme (mit Auswirkung auf Rechte und Pflichten von Menschen),
 - die Pflicht der Behörden, die Systeme im Verzeichnis einzutragen.
- Konkretisierung der einzelnen Transparenzelemente auf Verordnungsebene, z. B.:
 - Informationen zum verantwortlichen Organ,
 - die spezifischen Einsatzbereiche,
 - Informationen zur Art und Herkunft der bearbeiteten Daten und zu den Metadaten.

Weiterführende Informationen: *Nadja Braun Binder/Liliane Obrecht, Transparenz über den staatlichen Einsatz algorithmischer Entscheidungssysteme – Rechtliche Erwägungen ausgehend von Forderungen nach öffentlichen Verzeichnissen, AJP 10/2024, S. 1069 ff.*

Transparenz durch Begründung von Verfügungen

➤ Problemstellung

- Transparenz mit Blick auf konkrete Einzelfälle, d. h. Nachvollziehbarkeit von Verfügungen
- Begründungspflicht als Teilaspekt des rechtlichen Gehörs



Transparenz durch Begründung von Verfügungen

➤ **Anspruch auf inhaltlich und umfangmässig angemessene Begründung**

- Intensität des Eingriffs
- Komplexität der Sach- und Rechtslage
- Entscheidungsspielraum der Behörde
- Abweichung von einer behördlichen Praxis

➤ **Funktionen der Begründung**

- Rechtsmittelfunktion
- Akzeptanzfunktion
- Rechtssicherheitsfunktion
- Selbstkontrollfunktion
- Fremdkontrollfunktion

Transparenz durch Begründung von Verfügungen

➤ Auswirkungen auf die Funktionen der Begründung

- **Fehlende menschliche Beteiligung im Entscheidungsverfahren** als Herausforderung für
 - die Selbstkontrollfunktion (Stichwort: Automation Bias!)
 - die Akzeptanzfunktion
- **Fehlende Vorhersehbarkeit** als Herausforderung für
 - die Rechtsmittelfunktion
 - die Rechtssicherheitsfunktion
 - die Fremdkontrollfunktion
- **Fehlende Nachvollziehbarkeit der Information** als Herausforderung für
 - die Selbstkontrollfunktion
 - die Rechtsmittelfunktion
 - die Akzeptanzfunktion
 - die Fremdkontrollfunktion

Transparenz durch Begründung von Verfügungen

➤ **Teilweise geeignete Lösungsansätze**

- Kennzeichnungspflicht
- Weiterführende Informationen (z. B. zur Folgenabschätzung)
- Wahl geeigneter Modelle
- Menschliche Verantwortung

➤ **Zur Reduktion der Begründungsdichte**

- Grundsätzlich nicht gerechtfertigt
- Aber: ggf. Paradigmenwechsel durch retrospektive Systemkontrolle und Analyse der Entscheidungspraxis

Weiterführende Informationen: *Nadja Braun Binder/Liliane Obrecht, Die Begründung von Verfügungen beim Einsatz algorithmischer Systeme, SJZ 120/2024, S. 707 ff.*

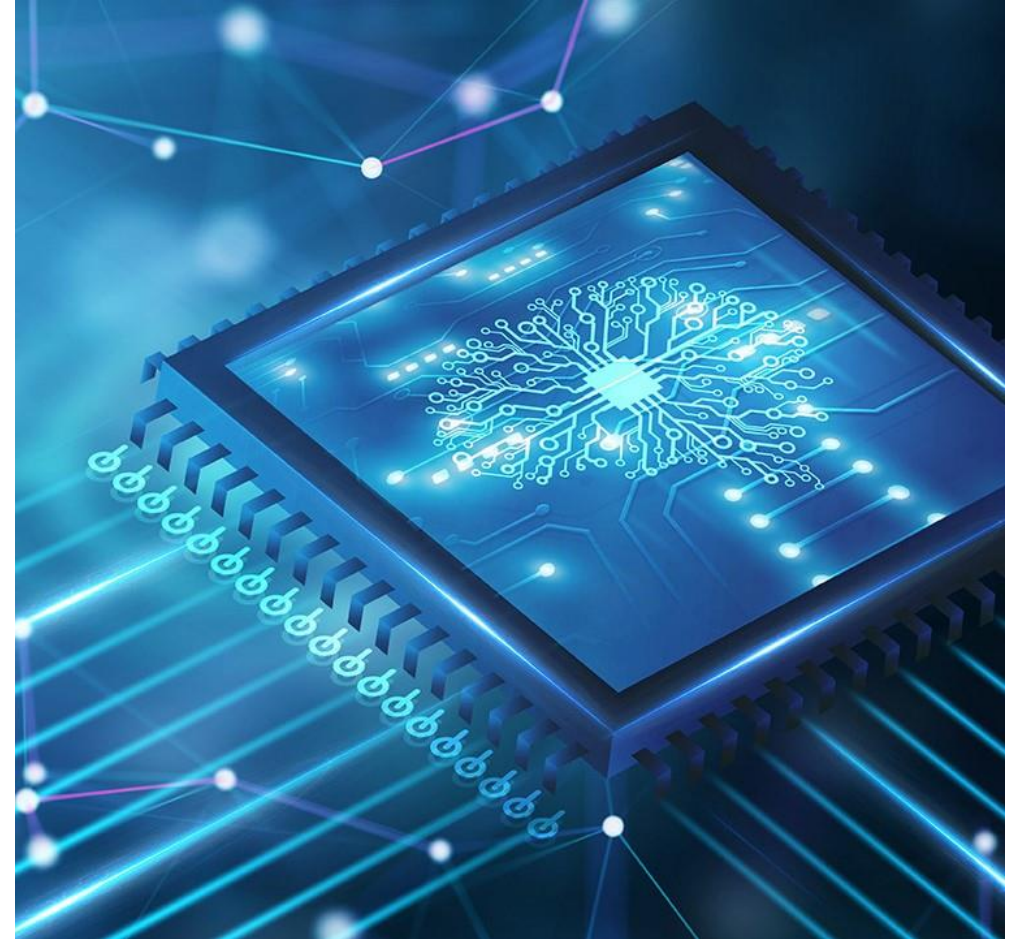
Transparenz

➤ Problemstellung

- Erkennbarkeit
- Erklärbarkeit / Nachvollziehbarkeit
- Beweisschwierigkeiten

➤ Lösungsansätze

- Informationspflicht bei Interaktion mit KI-System
- Informationspflicht bei voll- und teilautomatisierten Einzelentscheidungen
- Keine allgemeine Kennzeichnungspflicht für KI-generierte Inhalte, ev. sektorspezifische Vorgaben, bspw. für Medien
- Anwendung Prozessrecht, insb. Editionspflicht, sektorspezifische Dokumentationspflichten, punktuelle Beweislastumkehr



Diskriminierung

➤ Problemstellung

- Diskriminierungsverbot als Grundrecht bindet nur öffentliche Organe
- Fragmentarische Diskriminierungsverbote für Private: Gleichstellungsgesetz und Behindertengleichstellungsgesetz

➤ Lösungsansätze

- Allgemeines Gleichbehandlungsgesetz
- Verbot von Diskriminierung aufgrund der Bearbeitung der eigenen Personendaten: neuer Grundsatz der Datenbearbeitung
- Diskriminierung als Persönlichkeitsverletzung (Art. 28 ZGB)



Datenschutz

➤ Problemstellung

- Bearbeitung sehr grosser Mengen von Personendaten, insb. beim Training
- Transparenz, Zweckbindung, Datenminimierung und Speicherbegrenzung
- Verantwortlicher vs. Auftragsbearbeiter

➤ Lösungsansätze

- Anwendung des Datenschutzrechts
- Bearbeitung von Personendaten für Training als nicht-personenbezogene Bearbeitung
- Klärung der Verantwortlichkeit: Verantwortlicher vs. Auftragsbearbeiter



Manipulation

➤ Problemstellung

- «Grundangst»
 - Beeinträchtigung der Autonomie
 - Dominanz der Maschinen
- Schwer fassbares Phänomen

➤ Lösungsansätze

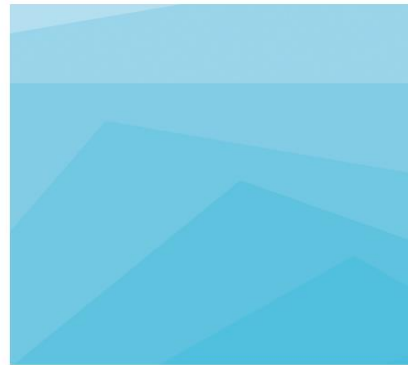
- Begriffsdefinition: Bewusstes Einwirken auf Entscheidung mit der Absicht, dass sich die betroffene Person selbst einen Nachteil zufügt
- Sektorspezifische Regelungen, bspw. «Bundesgesetz über Jugendschutz in den Bereichen Film und Videospiele»
- Wettbewerbsrecht (Art. 2 UWG)
- Allgemeines Persönlichkeitsrecht (Art. 28 ZGB)



Handlungsbedarf im Schweizer Recht

	Neue Gesetze	Anpassungen	Anwendung
Transparenz	✗	• Datenschutzgesetze	• Datenschutzgesetze
Datenschutz	✗	✗	• Datenschutzgesetze
Diskriminierung	• Gleichbehandlungsgesetz	• Datenschutzgesetze	• Art. 28 ZGB
Manipulation	✗	✗	• Art. 2 UWG • Art. 28 ZGB
Haftung	✗	• Produkthaftungsgesetz • Sektorspezifische Haftungsbestimmungen	• Art. 41 OR

Weitere Informationen



WHITE PAPER
**Transparenz durch
öffentliches Verzeichnis**
Nadja Braun Binder
Liliane Obrecht
Juni 2024



WHITE PAPER
**Transparenz
durch Begründung
von Verfügungen**
Nadja Braun Binder
Liliane Obrecht
Juni 2024



WHITE PAPER
Diskriminierung
Florent Thouvenin
Stephanie Volz
Juni 2024



WHITE PAPER
Datenschutz
Florent Thouvenin
Stephanie Volz
Juni 2024



WHITE PAPER
Transparenz
Florent Thouvenin
Stephanie Volz
Juli 2024



WHITE PAPER
Manipulation
Florent Thouvenin
Stephanie Volz
Juni 2024



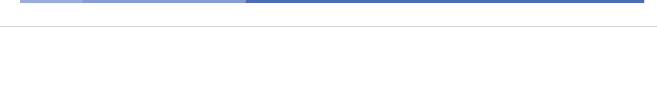
WHITE PAPER
Transparenz
Florent Thouvenin
Stephanie Volz
Juni 2024



WHITE PAPER
Diskriminierung
Florent Thouvenin
Stephanie Volz
Juni 2024



WHITE PAPER
Manipulation
Florent Thouvenin
Stephanie Volz
Juni 2024



- www.itsl.uzh.ch
- <https://ius.unibas.ch/de/e-piaf/>