

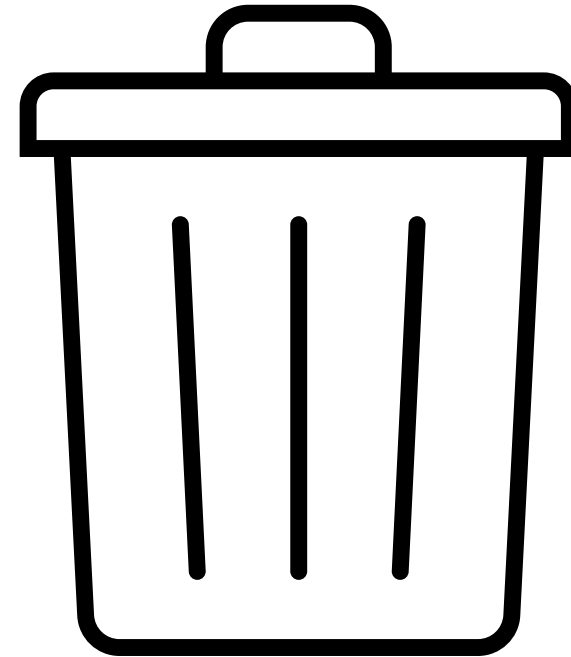


Responsive Rechtswissenschaft

Ringvorlesung an der Bucerius Law School
Herbsttrimester 25 | montags | 18 Uhr c.t.

Garbage in, liability out?

Möglichkeiten und Grenzen vertraglicher
Verantwortung für fehlerhafte KI-Trainingsdaten



Brussels, 19.2.2020
COM(2020) 65 final

WHITE PAPER

On Artificial Intelligence - A European approach to excellence and trust

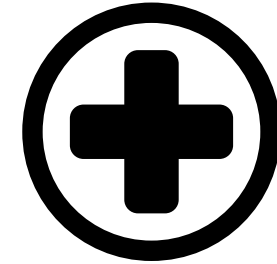
As discussed earlier, without data, there is no AI. The functioning of many AI systems, and the actions and decisions to which they may lead, very much depend on the data set on which the systems have been trained. The necessary measures should therefore be taken to ensure that, where it comes to the data used to train AI systems, the EU's values and rules are respected, specifically in relation to safety and existing legislative rules for the protection of fundamental rights. The following requirements relating to the data set used to train AI systems could be envisaged:



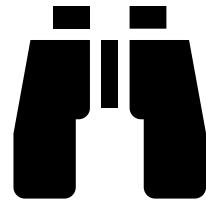
Bias in Recruiting



Bias in Credit Scoring



Bias in Health Care



Bias in Computer Vision



Bias in NLP



Agenda

- 1 **Datenqualität** als Grundbedingung maschinellen Lernens
- 2 **Verantwortlichkeit** für Datenqualität
- 3 **Status quo** des Rechtsrahmens für KI-Trainingsdaten
- 4 **Vertragliche Verantwortung** für „fehlerhafte“ KI-Trainingsdaten

Formen und Ziele

interdisziplinär arbeitender Rechtswissenschaft

1

Interdisziplinär informierte
Sachverhaltsaufarbeitung

3

Interdisziplinär informierte **Schärfung**
rechtlicher Begriffe und Konzepte

2

Interdisziplinär informierte
Maßstabsbildung

4

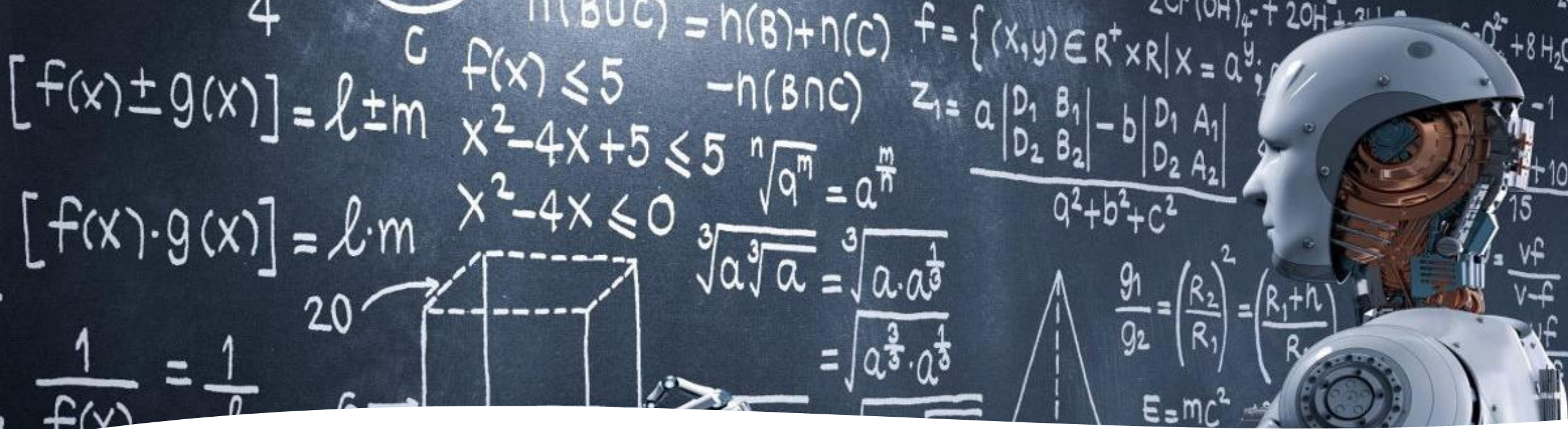
Interdisziplinär informierte
Identifikation von Forschungsbedarf



Interdisziplinär informierte
Sachverhaltsaufarbeitung

Datenqualität

als Grundbedingung maschinellen Lernens

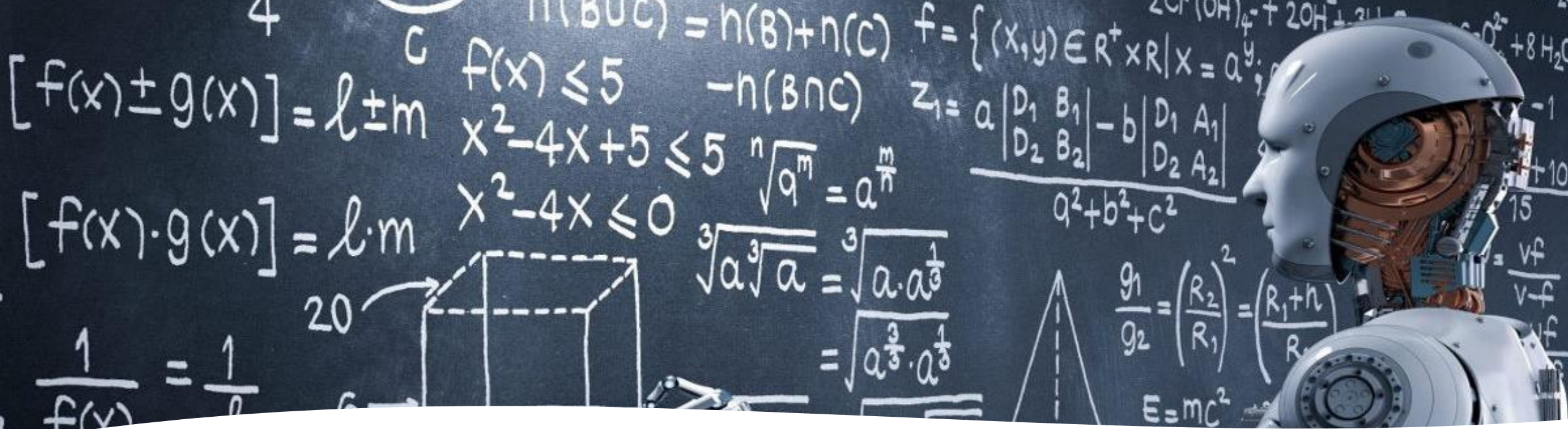


Machine learning



Grundprinzip

- Computer lernen aus Daten, statt nach festen Regeln programmiert zu werden
- Ziel: Erkennen von Mustern und Ableiten von Verallgemeinerungen

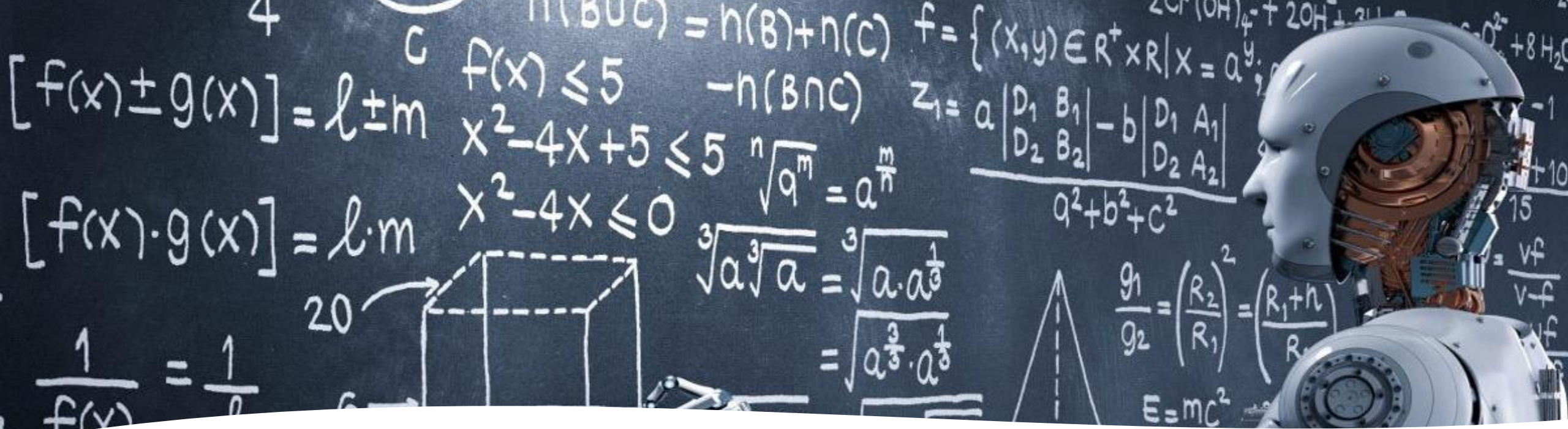


Machine learning



Funktionsweise

- In der Trainingsphase werden große Datenmengen verarbeitet
- Das Modell wird so angepasst, dass seine Vorhersagen möglichst genau mit den tatsächlichen Ergebnissen übereinstimmen



Machine learning



Charakter und Grenzen

- Machine learning-Modelle arbeiten probabilistisch, nicht deterministisch
- Sie berechnen Wahrscheinlichkeiten, keine sicheren Aussagen
- Die Güte der Vorhersage hängt unmittelbar von den Trainingsdaten ab:



„Garbage in, garbage out“



Interdisziplinär informierte **Maßstabsbildung**
und Kritik rechtlicher Entscheidungen

Maßstäbe

zur Ermittlung der Fehlerhaftigkeit von Trainingsdatensets

Qualitätsmaßstäbe



Abhängig von
Verwendungszweck

- ① Richtigkeit
- ② Aktualität
- ③ Faktordiversität und Vollständigkeit
- ④ Ausgewogenheit
- ⑤ Repräsentativität



Interdisziplinär informierte
Sachverhaltsaufarbeitung

Involvierte Akteure

im Rahmen der Wertschöpfung

Involvierte Akteure im Rahmen der Wertschöpfung



Datenlieferantinnen bzw. -quellen



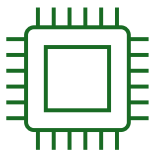
Plattformbetreiberinnen



Modellentwicklerinnen



Systemintegratorinnen



Hardware-Produzentinnen



Endnutzerinnen

Involvierte Akteure im Rahmen der Wertschöpfung



Datenlieferantinnen bzw. -quellen



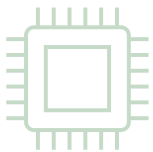
Plattformbetreiberinnen



Modellentwicklerinnen



Systemintegratorinnen



Hardware-Produzentinnen

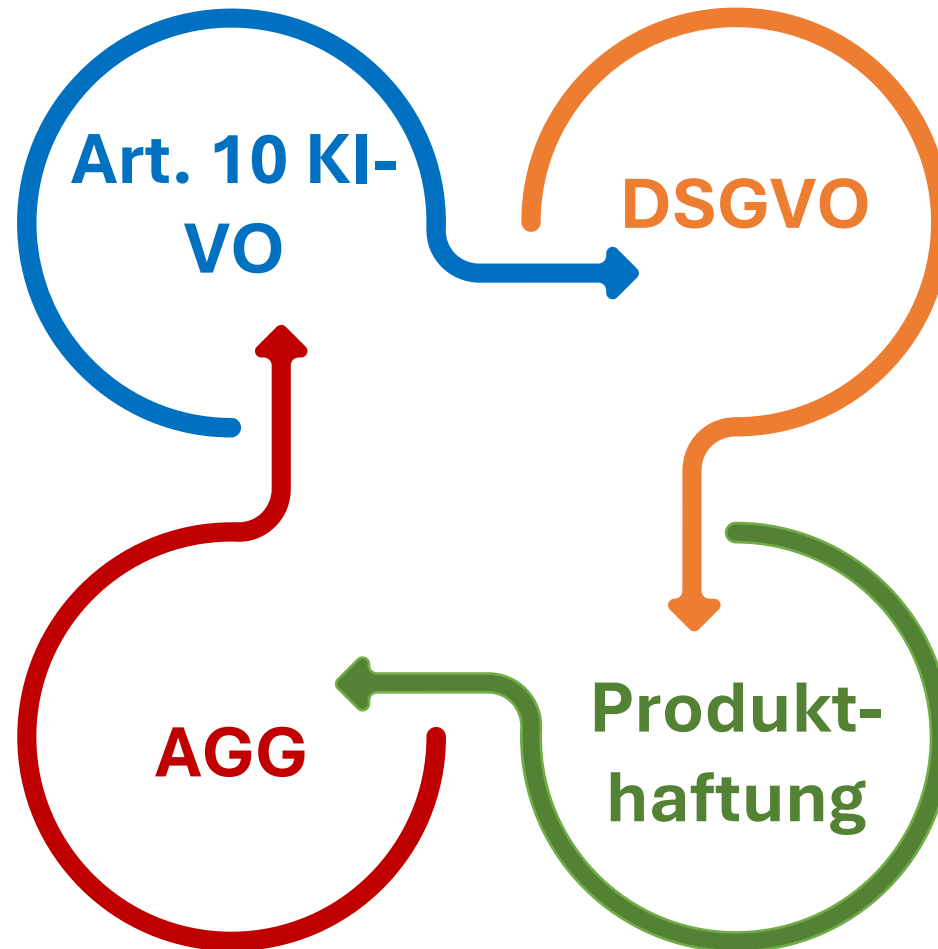


Endnutzerinnen

Status quo

des Rechtsrahmens für KI-Trainingsdaten

Status quo des Rechtsrahmens für KI-Trainingsdaten



Artikel 10

Daten und Daten-Governance

(1) Hochrisiko-KI-Systeme, in denen Techniken eingesetzt werden, bei denen KI-Modelle mit Daten trainiert werden, müssen mit Trainings-, Validierungs- und Testdatensätzen entwickelt werden, die den in den Absätzen 2 bis 5 genannten Qualitätskriterien entsprechen, wenn solche Datensätze verwendet werden.

(2) Für Trainings-, Validierungs- und Testdatensätze gelten Daten-Governance- und Datenverwaltungsverfahren, die für die Zweckbestimmung des Hochrisiko-KI-Systems geeignet sind. Diese Verfahren betreffen insbesondere

[...]

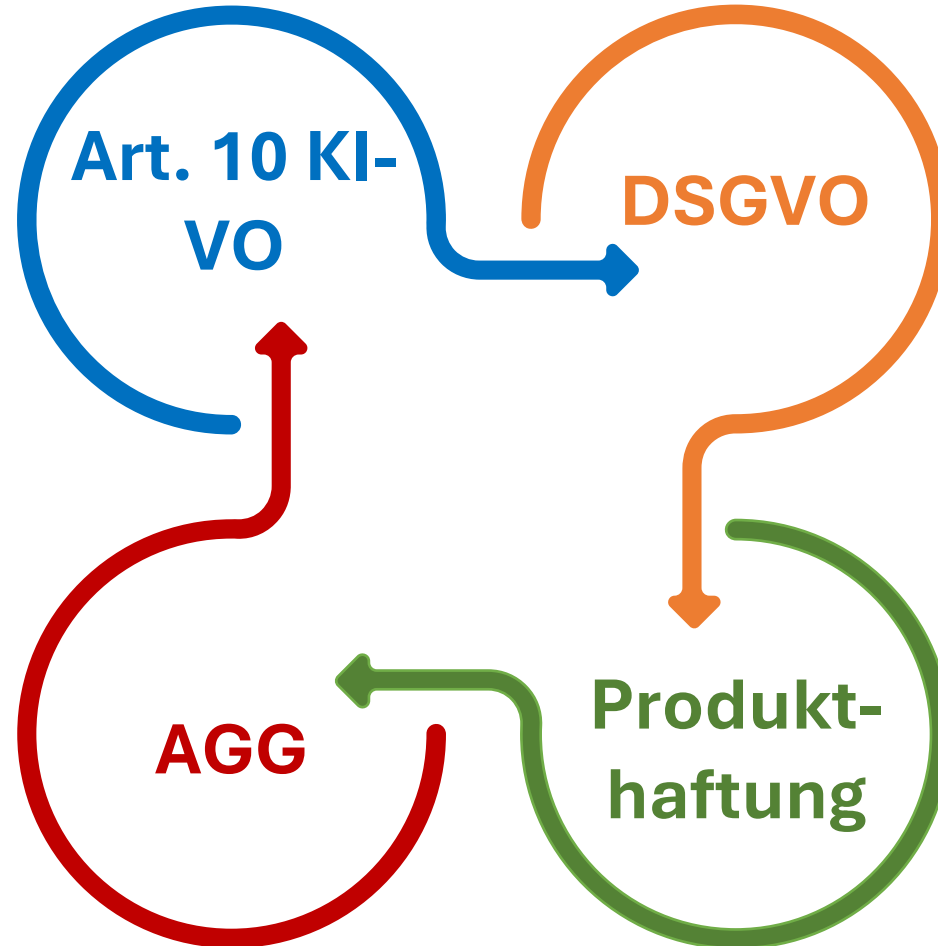


Ausgewogenheit/
Repräsentativität

- f) eine Untersuchung im Hinblick auf mögliche Verzerrungen (Bias), die die Gesundheit und Sicherheit von Personen beeinträchtigen, sich negativ auf die Grundrechte auswirken oder zu einer nach den Rechtsvorschriften der Union verbotenen Diskriminierung führen könnten, insbesondere wenn die Datenausgaben die Eingaben für künftige Operationen beeinflussen,
- g) geeignete Maßnahmen zur Erkennung, Verhinderung und Abschwächung möglicher gemäß Buchstabe f ermittelter Verzerrungen,

Status quo des Rechtsrahmens für KI-Trainingsdaten

- Nur Hochrisiko KI-Systeme
- Öffentlich-rechtliche Vorschrift
- Adressatin: Anbieterin iSd Art. 3 Nr. 3 KI-VO



KAPITEL II

Grundsätze

Artikel 5

Grundsätze für die Verarbeitung personenbezogener Daten

(1) Personenbezogene Daten müssen

[...]



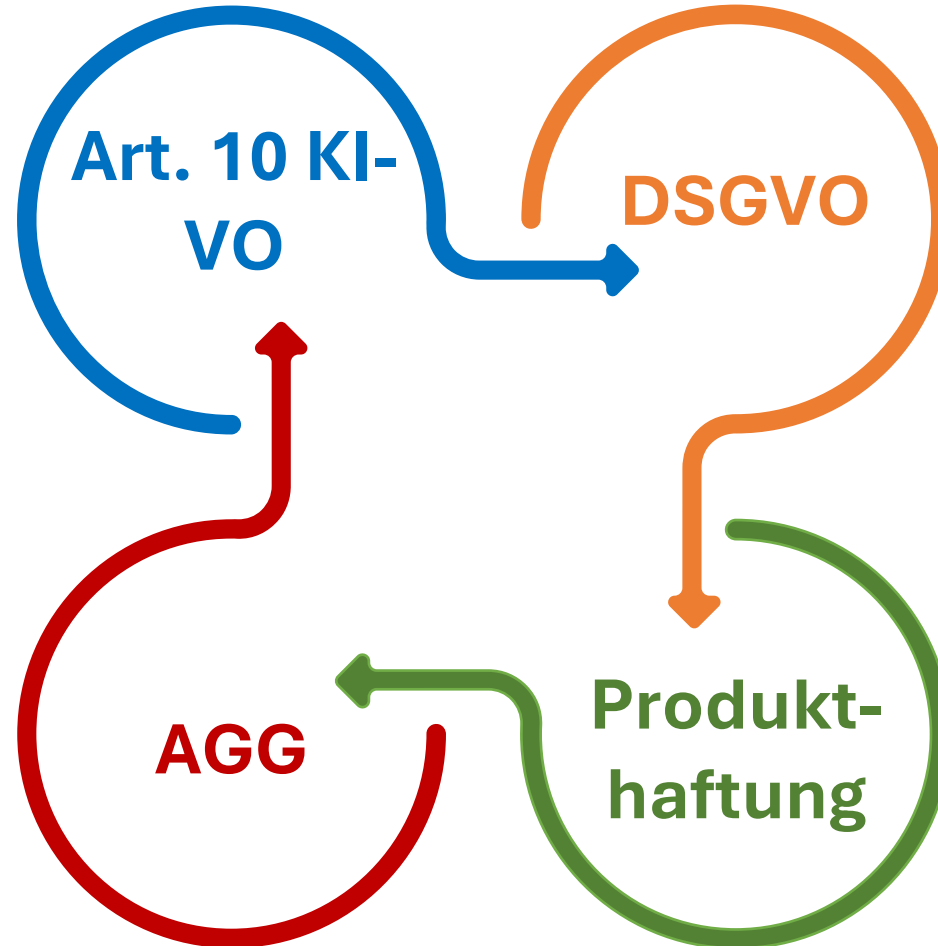
Richtigkeit und
Aktualität

- d) sachlich richtig und erforderlichenfalls auf dem neuesten Stand sein; es sind alle angemessenen Maßnahmen zu treffen, damit personenbezogene Daten, die im Hinblick auf die Zwecke ihrer Verarbeitung unrichtig sind, unverzüglich gelöscht oder berichtigt werden („Richtigkeit“);

Status quo des Rechtsrahmens für KI-Trainingsdaten

- Nur Hochrisiko KI-Systeme
- Öffentlich-rechtliche Vorschrift
- Adressatin: Anbieterin iSd Art. 3 Nr. 3 KI-VO

- Allenfalls diskriminierungs-spezifische Rechtsfolgen fehlerhafter KI-Trainingsdaten



- Nur einzelne Qualitätsmerkmale adressiert
- Personenbezug zunehmend nicht vorhanden

- ProdHaftG: Haftung nur für körperliche Gegenstände
- Beschränkte Schutzgüter

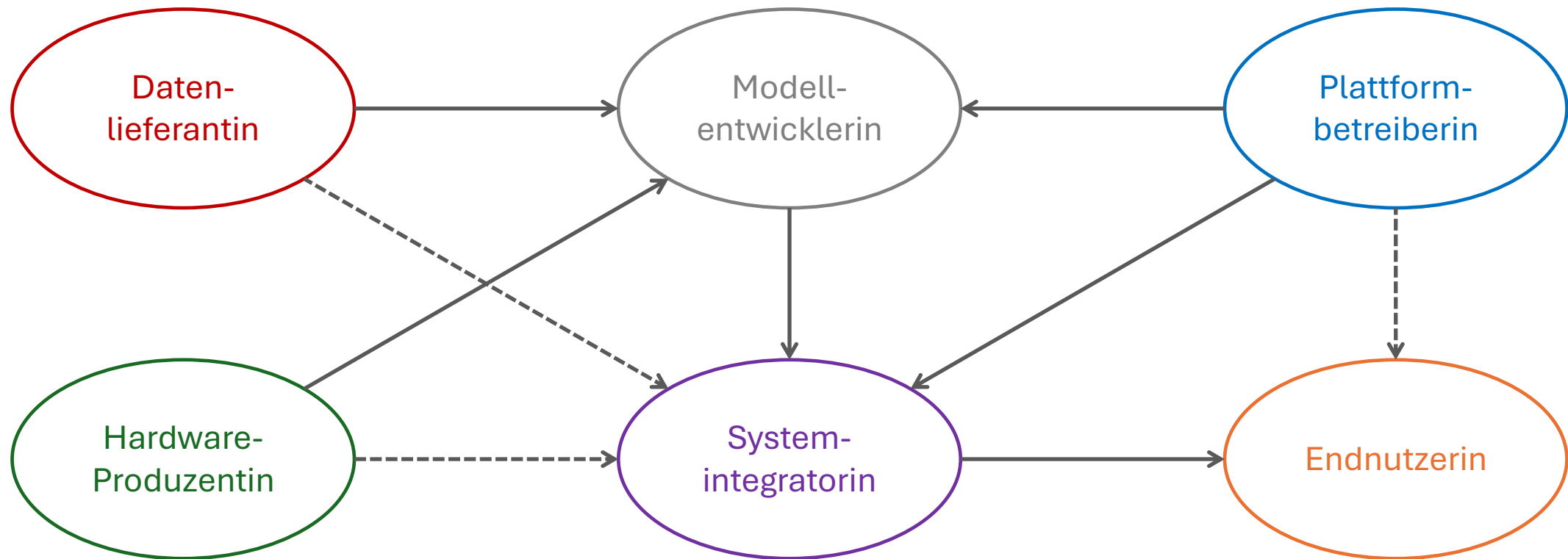


Interdisziplinär informierte **Schärfung**
rechtlicher Begriffe und Konzepte

Vertragliche Verantwortung

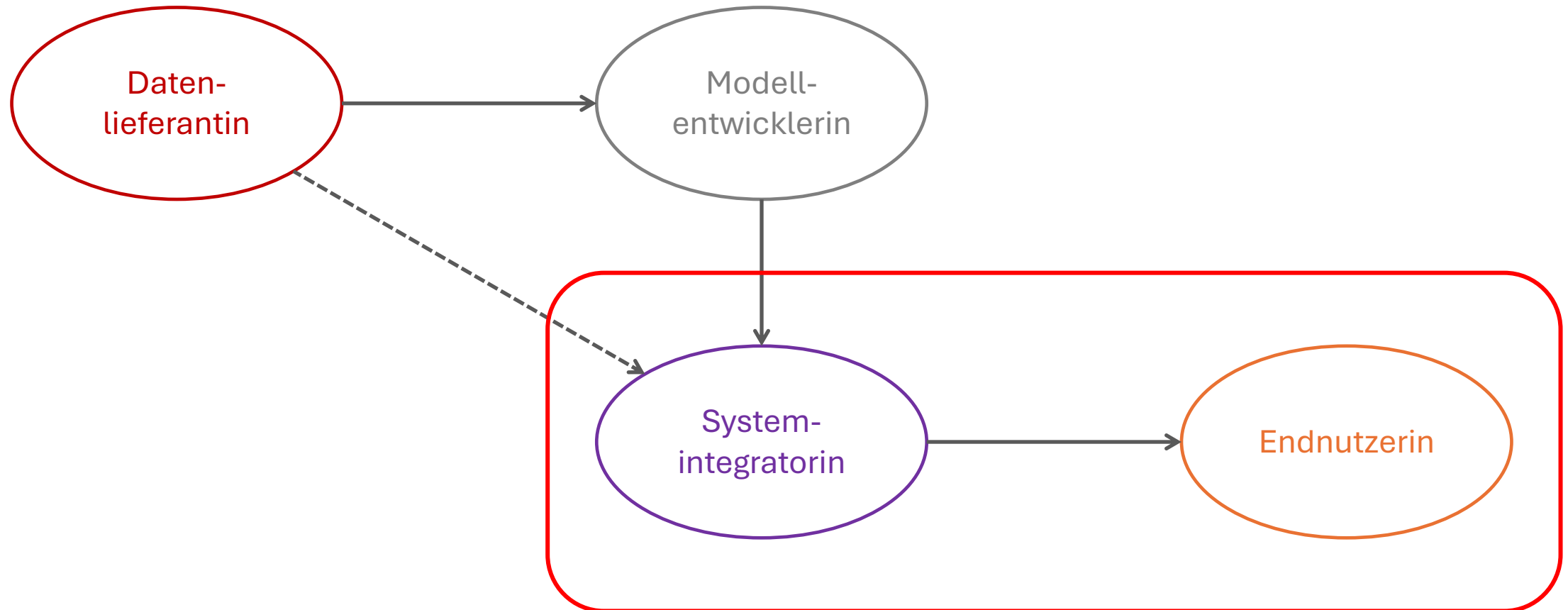
für „fehlerhafte“ KI-Trainingsdaten

Vertragliche Verantwortung für „fehlerhafte“ KI-Trainingsdaten

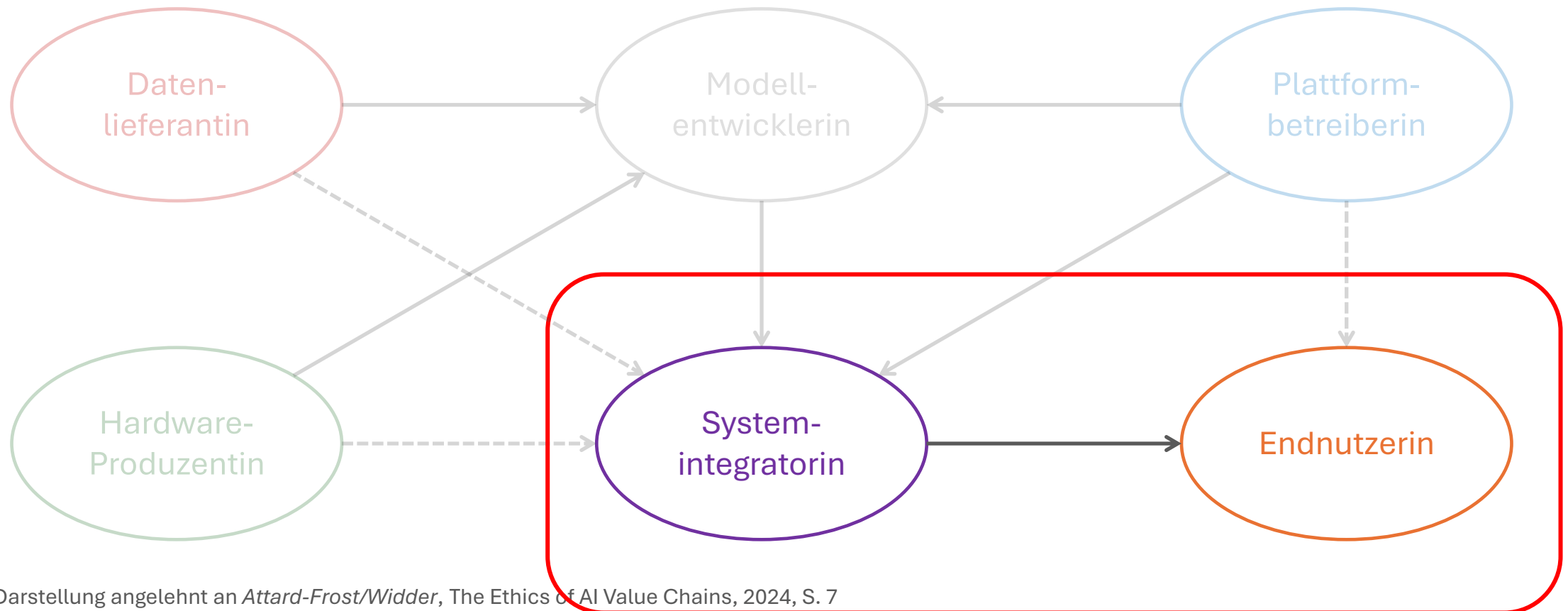


Darstellung angelehnt an Attard-Frost/Widder, The Ethics of AI Value Chains, 2024, S. 7

Vertragliche Verantwortung für „fehlerhafte“ KI-Trainingsdaten

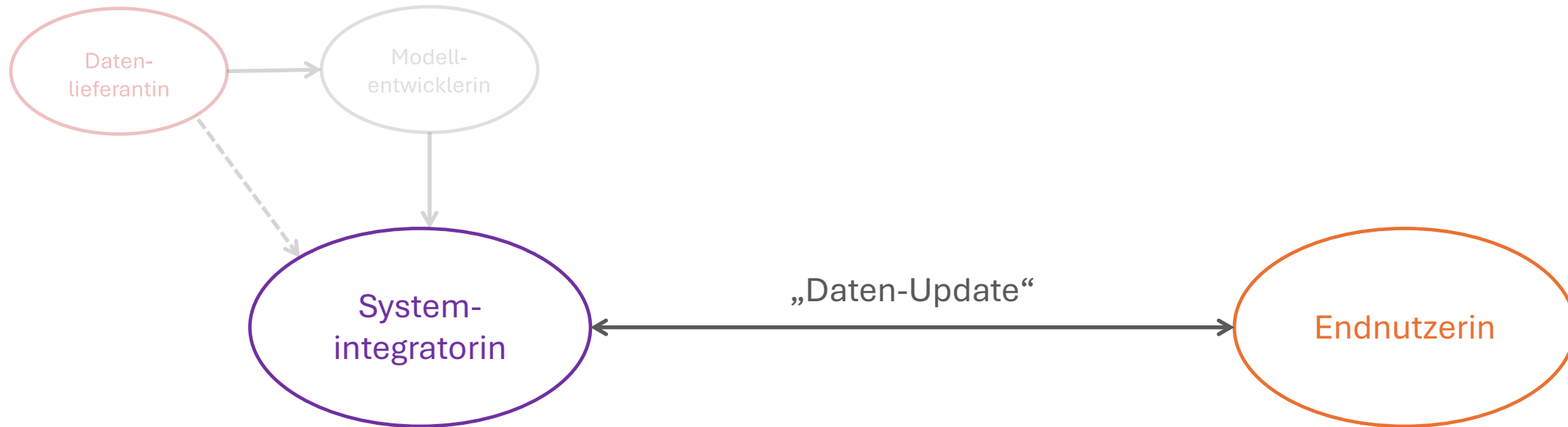


Vertragliche Verantwortung für „fehlerhafte“ KI-Trainingsdaten

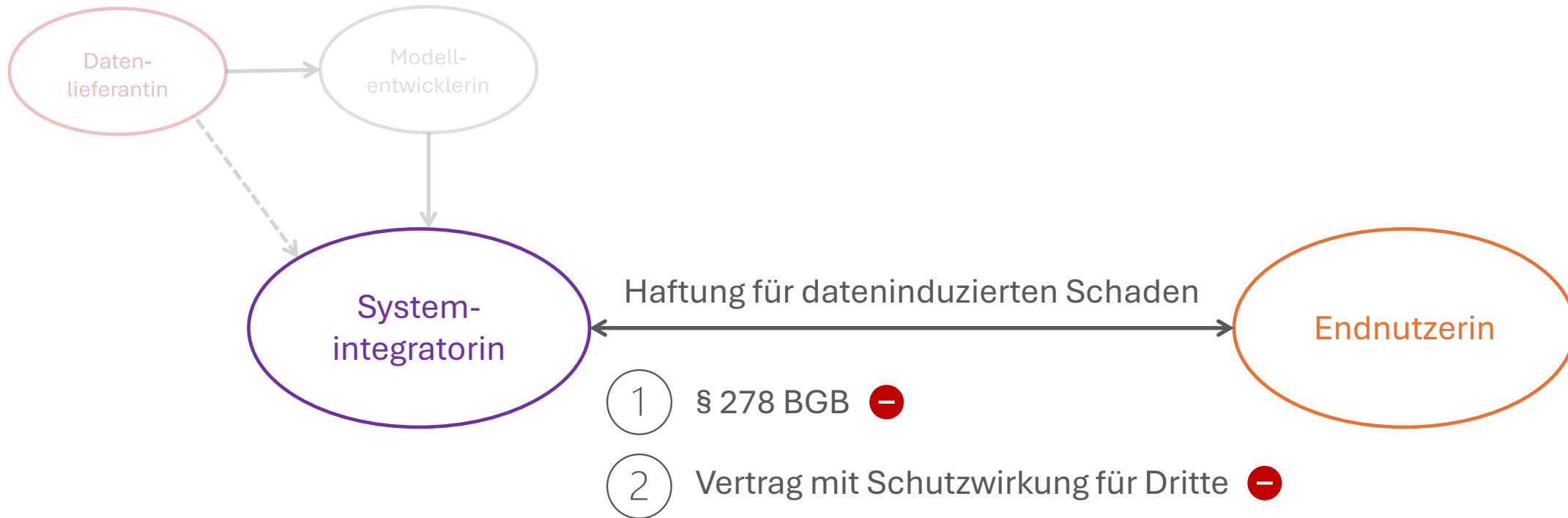


Darstellung angelehnt an Attard-Frost/Widder, The Ethics of AI Value Chains, 2024, S. 7

Vertragliche Verantwortung für „fehlerhafte“ KI-Trainingsdaten



Vertragliche Verantwortung für „fehlerhafte“ KI-Trainingsdaten





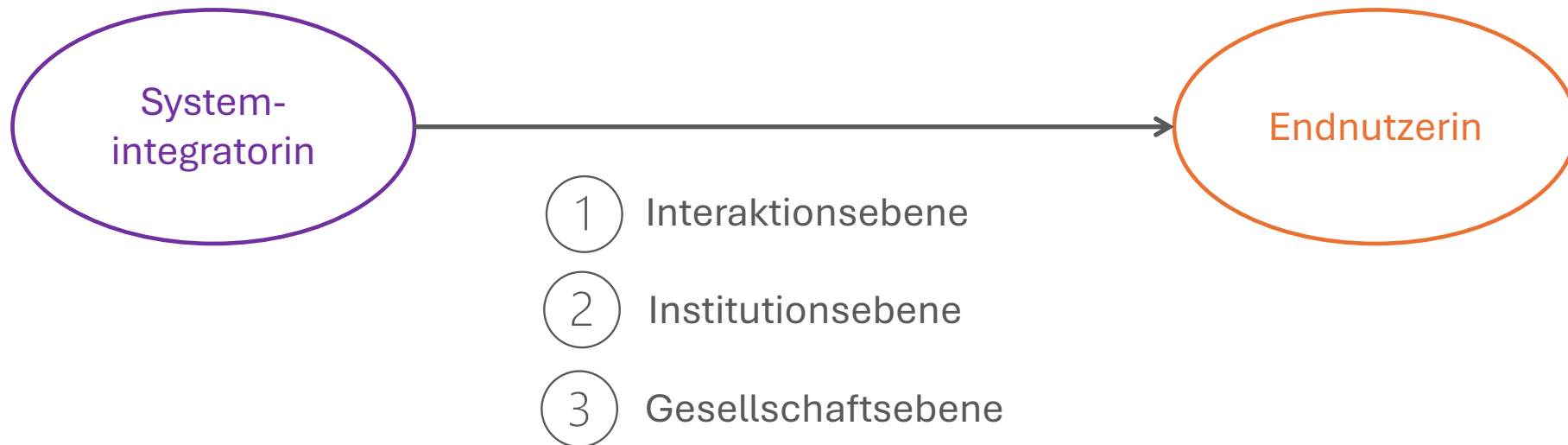
Interdisziplinär informierte **Schärfung
rechtlicher Begriffe und Konzepte**

Materialisierung des Vertragsrechts

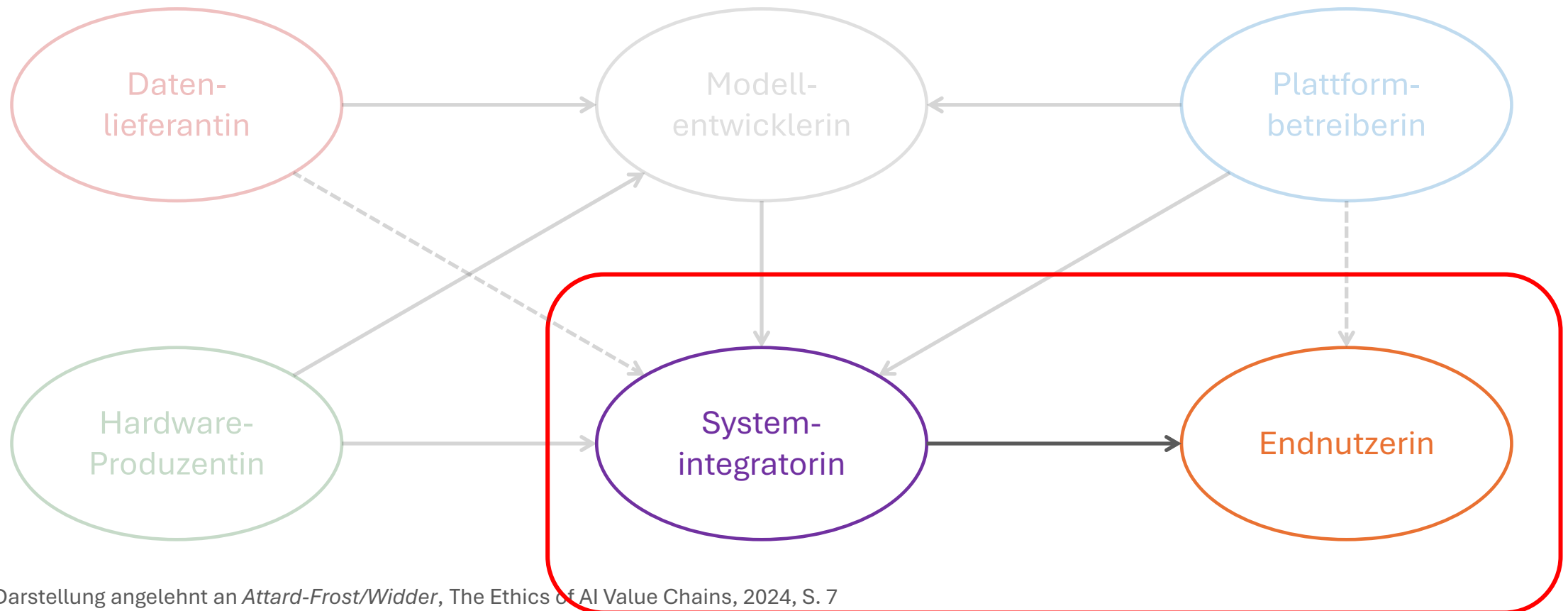
zur Rekonstruktion der Vertragsbeziehungen



Materialisierung des Vertragsrechts

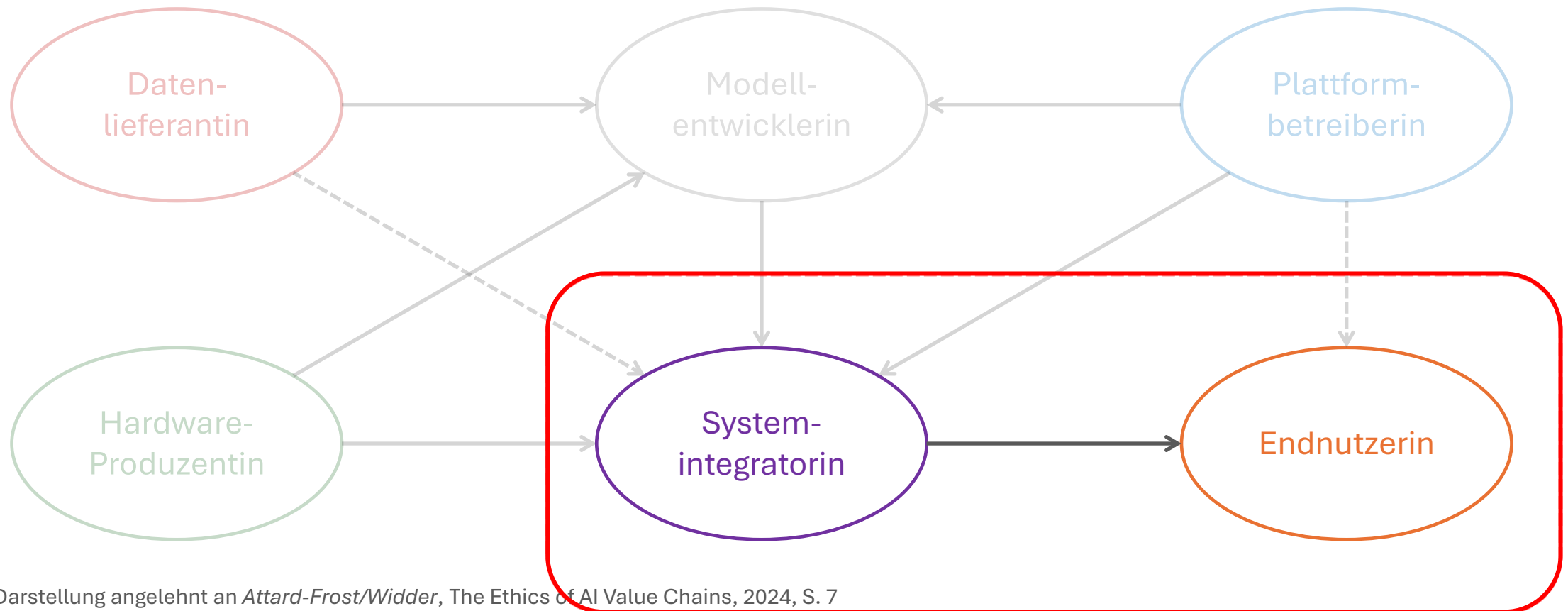


Materialisierung des Vertragsrechts



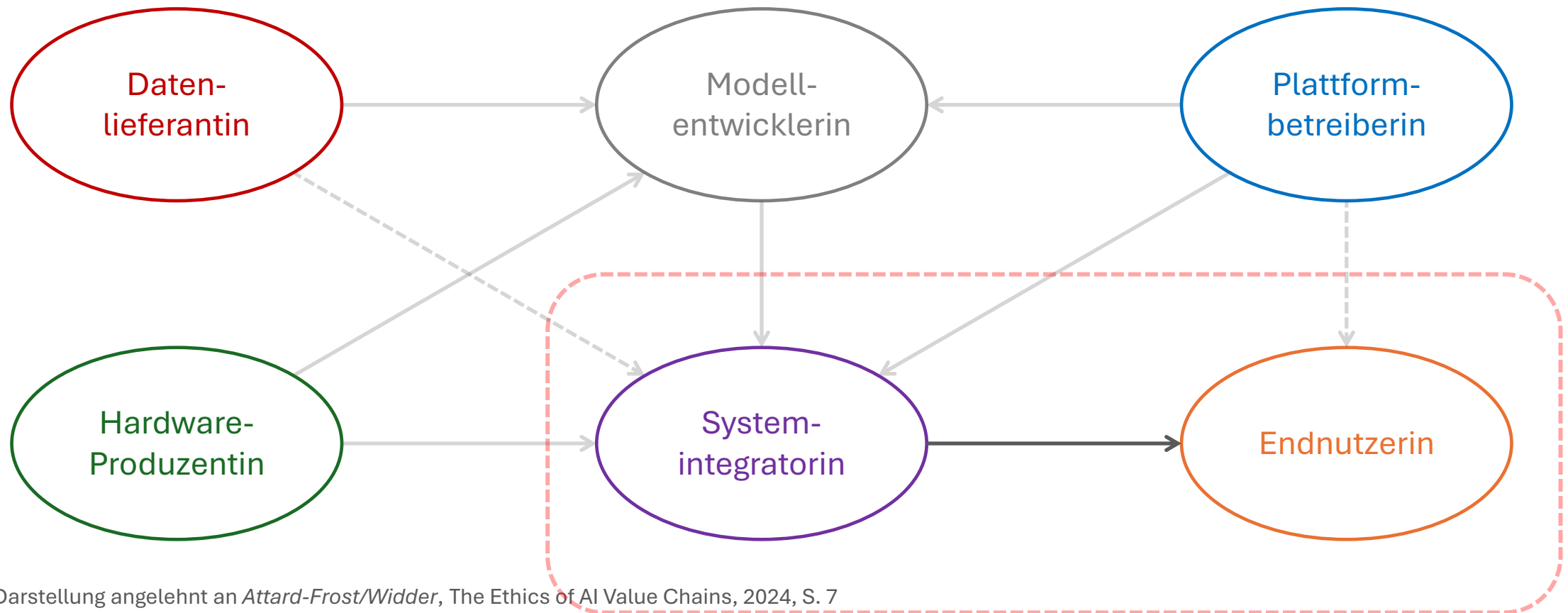
Darstellung angelehnt an Attard-Frost/Widder, The Ethics of AI Value Chains, 2024, S. 7

Materialisierung des Vertragsrechts



Darstellung angelehnt an Attard-Frost/Widder, The Ethics of AI Value Chains, 2024, S. 7

Materialisierung des Vertragsrechts

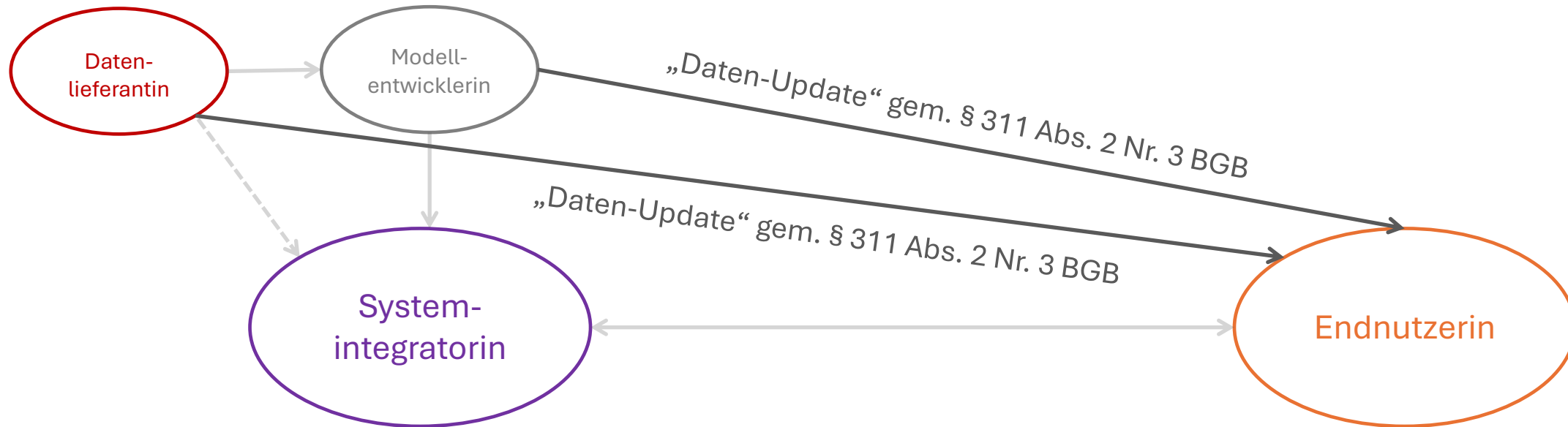


Darstellung angelehnt an Attard-Frost/Widder, The Ethics of AI Value Chains, 2024, S. 7

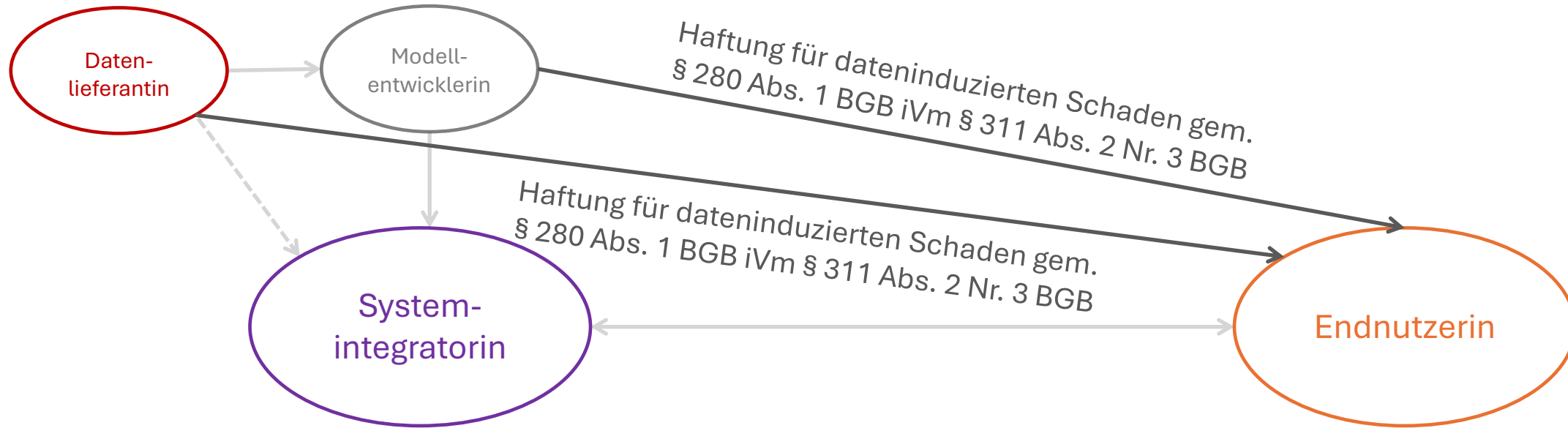
Vertragsnetze

im System des BGB

Vertragsnetze im System des BGB



Vertragsnetze im System des BGB



Gegenüberstellung

der vorgestellten Lösungen

Gegenüberstellung der gefundenen Lösungen

Perspektive des klassischen Vertragsrechts



Reduktion von Komplexität



Im Einklang mit tradierten Prinzipien



Regulierungsfunktion des Vertragsrechts?

Materialisierung des Vertragsrechts



Überzeugungskraft der gefundenen Ergebnisse



Relativitätsdogma seit jeher Frage zeitgenössischer Wertungen



Vertragsrecht flexibel genug, um dynamischer Entwicklung zu folgen

Kernaussagen

- Verträge müssen nicht notwendig als bipolare Austauschbeziehung verstanden werden
- Auf der Grundlage eines erweiterten vertragstheoretischen Verständnisses lassen sich auch netzwerkartige Wertschöpfungsprozesse vertraglich erfassen
- Eine derartige Materialisierung des Vertragsrechts eröffnet insbesondere die Möglichkeit, die Verantwortung für fehlerhafte Trainingsdaten bei derjenigen Marktteilnehmerin zu verorten, die tatsächlichen Einfluss auf die Datenqualität hat
- Dadurch entstehen ökonomische Anreize, qualitativ hochwertige Datensätze bereitzustellen und zu nutzen

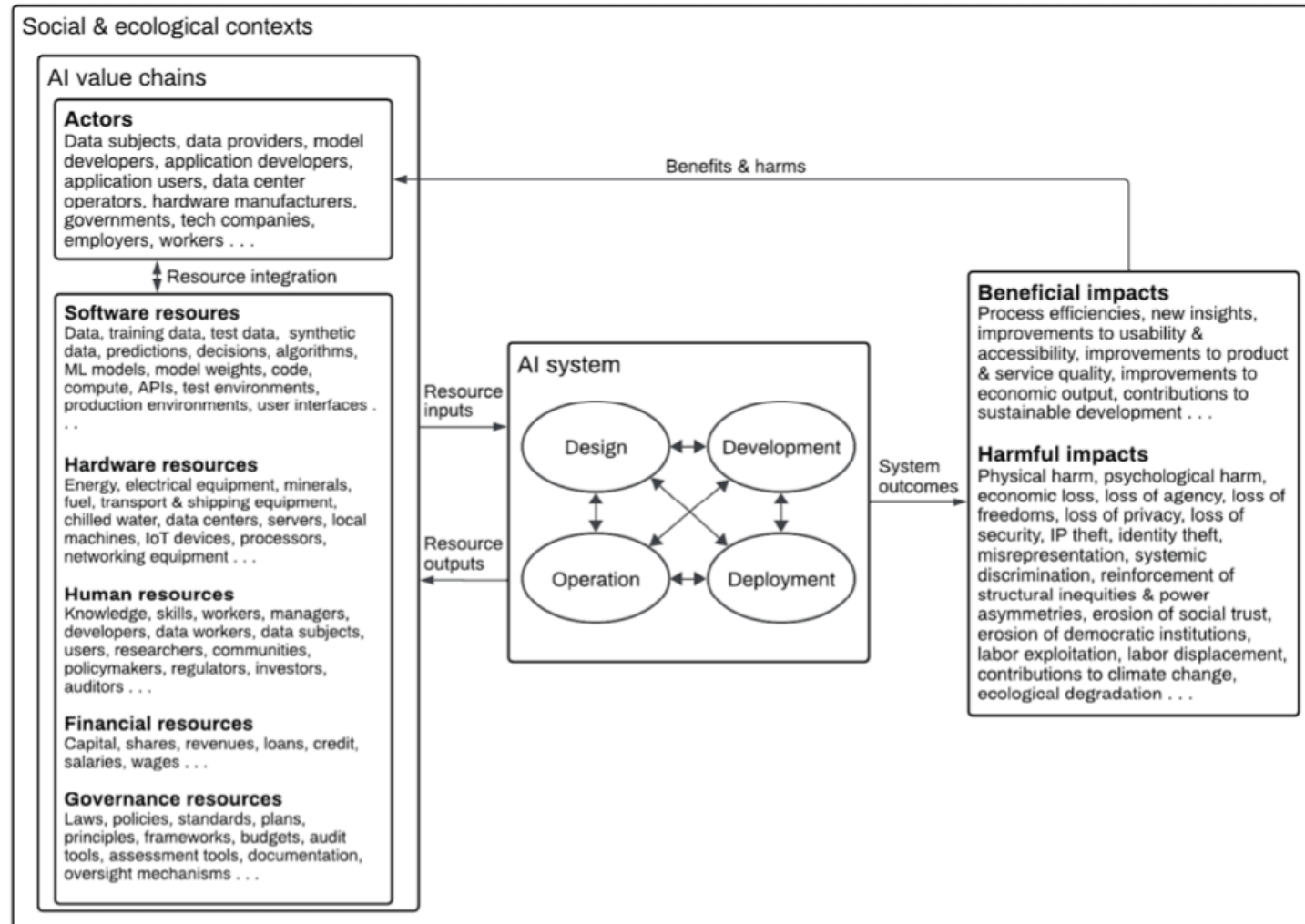




Interdisziplinär informierte **Identifikation** von **Forschungsbedarf**



AI Ecosystem



Vielen Dank!

katharina.wunner@uni-bayreuth.de