IAAI
Institute for Applied Artificial Intelligence

Rechtlicher Schutz von Daten in KI-Forschungs- und Entwicklungsprojekten

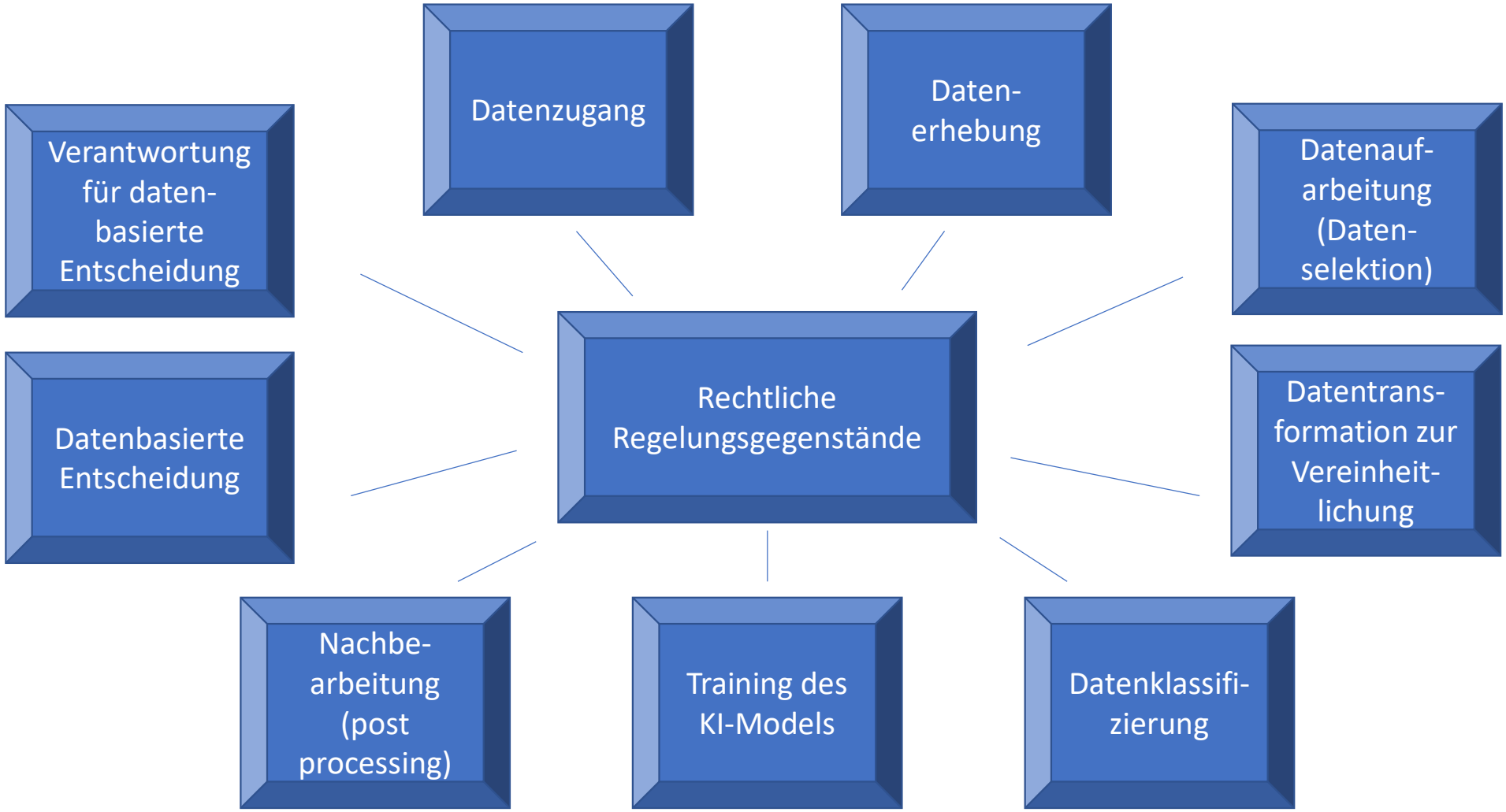
Shutterstock/Evannovostro

I. Zum Begriff der „Daten“ und zu den Funktionsprinzipien der Daten-nutzung

KI-Trainingsdaten als Grundlage von KI-Systemen

1. Einsatz von Trainingsdaten und Testdaten im Bereich des maschinellen Lernens

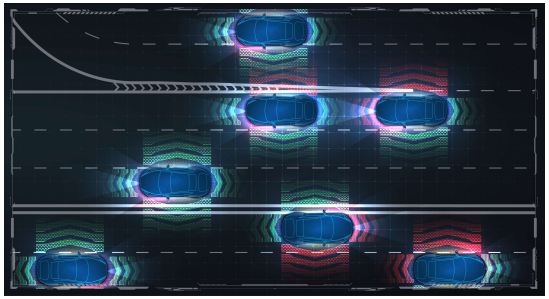
2. Aufwand bei der Erstellung von KI-Trainingsdaten
   (1) Ermittlung geeigneter Daten
   (2) Bestimmung einheitlicher Datenformate
   (3) Identifizierung fehlender Werte, Ausreißer und Ergänzung von Daten
   (4) Kontrolle und Korrektur durch mögliche Experten auf dem Gebiet des zu trainierenden Modells
   (5) Überprüfung der Daten auf Grundlage des Datenschutzes, der Diskriminierungsfreiheit etc.


Datenzugang
Daten-
erhebung
Verantwortung
für daten-
basierte
Entscheidung
Datenauf-
arbeitung
(Daten-
selektion)
Rechtliche
Regelungsgegenstände
Datenbasierte
Entscheidung
Datentrans-
formation zur
Vereinheit-
lichung
Nachbe-
arbeitung
(post
processing)
Training des
KI-Models
Datenklassifi-
zierung

Dilemma der fehlenden KI-Trainingsdaten

1. Befund in vielen KI-Projekten und Anwendungsgebieten:
   - 1.1 Geringe Datenquantität: keine robusten Modelle
   - 1.2 Geringe Datenqualität (unvollständig, veraltet, inkonsistent, unausgewogen etc.): fehlerhafte KI System

2. Beispiele



Shutterstock/ZinetroN

Shutterstock/Coffeemill

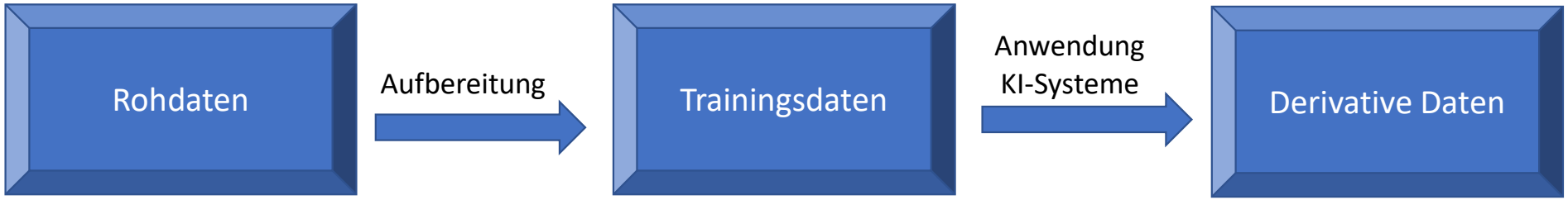
Shutterstock/asharkyu

Rechtliche Regelungsaufgaben

1. Rechtliche Anreizmechanismen
   - 1.1 Anreizmechanismen zur Erstellung von Daten = Schaffung eines Marktes für Trainingsdaten
   - 1.2 Anreizmechanismen zur Gewährung des Zugangs zu bestehenden Daten

2. Anreize im Rahmen der rechtlichen Zugangsregelungen
   - 2.1 Schutz der Dateninhalte (z. B. Know-How, personenbezogene Daten etc.)
   - 2.2 Schutz der Leistungen im Rahmen der Aufbereitung von Daten
   - 2.3 Access Rights zur Schaffung von Nutzungsmöglichkeiten
     - aber: Negative Lenkungswirkung von Zugangsansprüchen

Differenzierung des Datenbegriffs



- Einzeldaten oder umstrukturierte Daten etc.
- Mögliche Kontextualisierung

- Strukturierte Datensets
- Vorbereitung für den Einsatz als Trainingsdaten

Differenzierung zwischen KI-Trainingsdaten und Rohdaten in der rechtlichen Behandlung

1. Beispiele für Rohdaten

   - Industrielle Rohdaten (Sensor- oder Betriebsdaten von Maschinen und Anlagen)
   - Nutzerdaten (im Bereich Wearables, Social Media oder Online-Handel)
   - Wissenschaftliche Daten (Versuchs- und Messdaten aus Experimenten in Laborversuchen, Beobachtungsdaten)
   - Staatliche Überwachungsdaten (Verkehrs- und Bewegungsdaten)
   - Medizinische Diagnosedaten

Differenzierung zwischen KI-Trainingsdaten und Rohdaten in der rechtlichen Behandlung

2. Mögliche Gründe für den rechtlichen Schutz von Rohdaten

   2.1 Schutz der Rohdateninhalte

   - Verkörperung von Geschäftsgeheimnissen (Logistikdaten, Maschinenbetriebsdaten, Forschungsdaten)
   - Personenbezogene Daten (Nutzerprofile, genetische Informationen)
   - Sicherheitsrelevante Daten
   - etc.

   2.2 Keine rechtlich schutzwürdige „Veredelungsleistung“ bei der Bereitstellung von Rohdaten?

3. Schlussfolgerung:

   Unterschiedliche rechtliche Schutzprofile bei Rohdaten und KI-Trainingsdaten

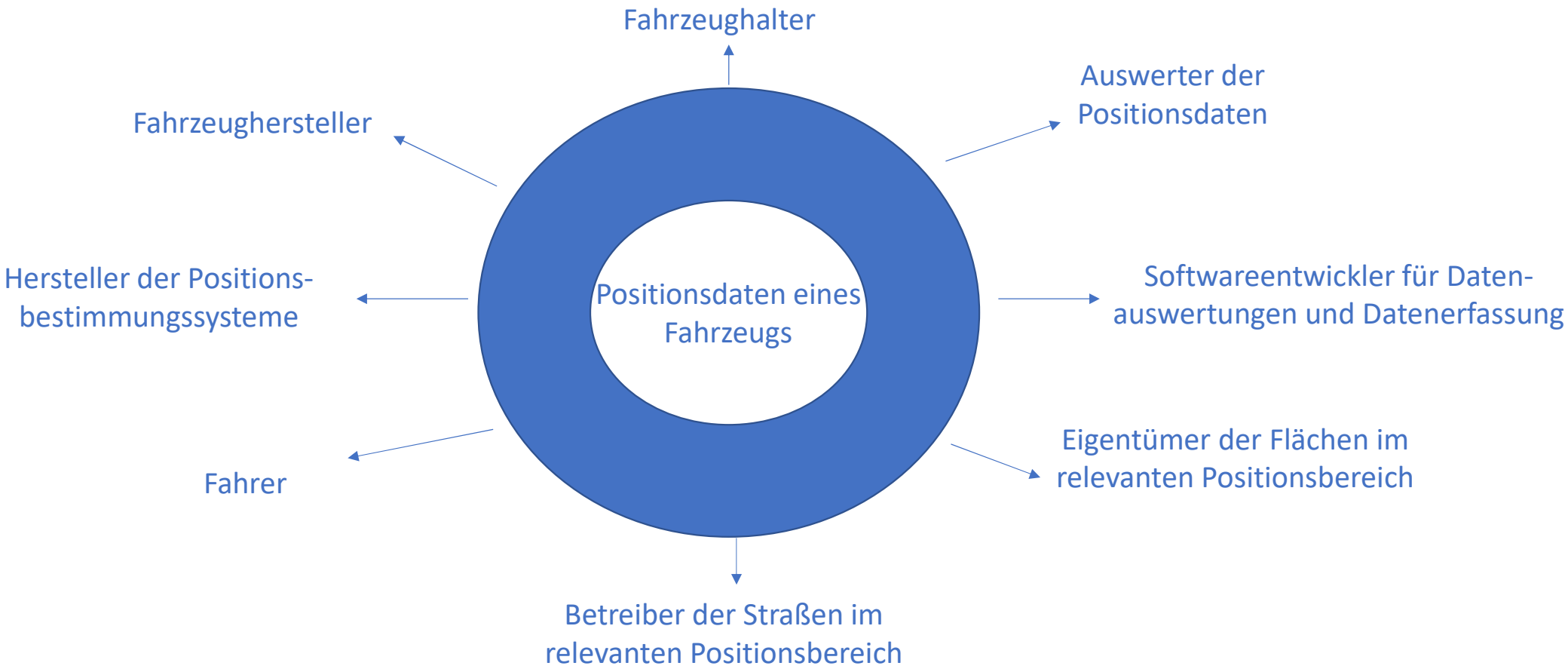
Folgen der Gewährung des rechtlichen Schutz für Daten

1. Gewährung von Monopolpositionen durch den rechtlichen Schutz von Daten als Belohnung für Datenerstellung und Zugangsgewährung

   1.1 Justierung des Umfang von Verbietungsrechten – Blockierungsrisiken?

   1.2 Inhaber der Rechte – Wer verdient die Belohnung?
   - Datenbereitsteller?
   - Datenveredler?
   - Inhaberposition in bilateralen oder multilateralen Entwicklungsprojekten?

Rechteinhaberschaft?



Sicherung durch den rechtlichen Schutz

2. Reduzierung von Kontrollverlustrisiken durch den rechtlichen Schutz

2.1 Daten als Grundlage von Wettbewerbspositionen und individuellen Rechten

2.2 Kontrollverlust bereits durch Daten*zugang*

- Risiko insbesondere bei der Verkörperung von geheimen Informationen in den Daten
- Lösung: Zugangsregulierung/“Need-to-Know“-Prinzip

2.3 Kontrollverlust durch fehlende Nutzungs*regelungen*

- Lösung: Nutzungsbedingungen/Lizenzen

Qualifizierung von Daten als handelbares Gut durch den rechtlichen Schutz

3. Rechtliche Kompensationsmodelle durch Gewährung von Gegenleistungen für Zugangs- und Nutzungsgewährungen

   3.1 Finanzielle Gegenleistung

   3.2 Andere Gegenleistungen

   - Datenbereitsteller/Datenveredler erhält Zugang zu KI-basierten Systemen bzw. KI-genierten Ergebnissen

Interessensausgleich in KI - F&E Projekten

1. Allgemeine Akteure
   - 1.1 Rohdatenzulieferketten
   - 1.2 „Datenveredlern“ für die Aufbereitung und/oder das Training von Trainingsdaten
   - 1.3 Entwickler der KI-Modelle
   - 1.4 Datenerfassungsanbietern (Inspektionssysteme, Mess- und Sensortechnik etc.)
   - 1.5 Sonstige Systemanbieter
     - Hohe Hard- und Softwareinfrastrukturanforderungen auf allen Entwicklungs- und Anwendungsebenen

Interessensausgleich in KI – F&E Projekten

2. Datennutzungen im Kooperationsmodell, z.B.

   2.1 Datenerfassung durch Nutzer mit KI basierter Auswertung im Nutzer- und/oder Herstellerinteresse

   - Wearables, Fahrzeugdaten, SmartHome, medizinische Diagnosesysteme etc.
   - Auswertung von Werkzeugnutzungsdaten in der Zusammenarbeit zwischen Werkzeughersteller sowie Betreiber der Produktion (Smart Factory)
   - Landwirte/Pflanzenschutzmittelhersteller beim SmartFarming

   2.2 Produkt- oder Verfahrensentwicklung durch KI basierte Optimierungsverfahren aufgrund interner oder externer Daten

   - Biotechnologische Forschung, Konstruktionssimulationen

Integrations- und Kooperationsdruck als Grundlage neuer rechtlicher Regelungsmodelle

3. Rechtliche Regelungsaufgaben in Kooperationen

3.1 Regelung der Datenfragen als Haupt- oder Nebenelemente einer Kooperation

- z.B. Verwertung der im Kooperationsprojekt generierten Daten durch Datennutzungsverträge mit Dritten

3.2 Streitvermeidung (z.B. Lufthansa (Aviation Data Hub) vs. Airbus zur Datenplattform (Skywise), die im Flugzeugbetrieb generierte Daten zur Funktion von Komponenten erfasst)

- Regelung der Rechtezuordnung und Rechtenutzung

II. Rechtlicher Schutz von Daten

Rechtlicher Schutz von Rohdaten und KI-Trainingsdaten

1. Kein Sacheigentum nach §§ 903, 90 BGB, da Beschränkung auf körperliche Gegenstände

2. Keine schutzrechtliche Sicherung (Patente etc.) für isolierte Daten

3. Urheberrechtlicher Schutz nach § 2 Abs. 1 Nr. 7 UrhG
   - nur bei persönlich geistiger Schöpfung nach § 2 Abs. 2 UrhG
   - wohl zumindest bei Rohdaten im Regelfall zu verneinen

4. Schutz als „Datenbankwerk“ nach § 4 Abs. 2 UrhG, wenn die Auswahl und Anordnung der Elemente auf einer schöpferischen Leistung beruhen
   - bei Trainingsdaten kann in Abhängigkeit von der Komplexität ein Schutz gegeben sein

5. Schutz als Computerprogramm nach § 69a UrhG
   - bei Daten zu verneinen

Schutz des Datenbankherstellers nach § 87a UrhG

1. Schutz von Datenbanken nach § 87a UrhG

   1.1 Sammlung von Werken, Daten oder anderen unabhängigen Elementen, „die systematisch oder methodisch angeordnet ...“ sind

   1.2 „deren Beschaffung, Überprüfung und Darstellung eine nach Art und Umfang wesentliche Investition erfordert“

2. Kriterium einer „Datenbank“ dürfte bei der Sammlung von Rohdaten oder Trainingsdaten, die über Einzeldaten hinausgehen, erfüllt sein

Schutz des Datenbankherstellers nach § 87a UrhG

3. Kriterium „wesentlichen Investition“ (strittig)
   - Herrschende Meinung: Investitionen in die *Erzeugung* von Daten, das heißt Kosten für Maschinen, Sensoren, Messsysteme etc., werden nicht erfasst, sondern nur Kosten in die Beschaffung und Strukturierung

4. *Im Regelfall dürfte ein Schutz als Datenbank nach § 87a UrhG zu bejahen sein, soweit nicht allein Einzeldaten oder gänzlich unstrukturierte Datensets betroffen sind*

Wettbewerbsrechtlicher Leistungsschutz nach § 4 Nr. 3 UWG

1. Daten müssen eine „wettbewerbliche Eigenart“ aufweisen und es müssen Unlauterkeitsumstände hinzutreten

2. In Bezug auf Daten ergibt sich eine praktische Relevanz nur in Fällen einer unberechtigten Umgehung von Schutzmaßnahmen (ehemalige Mitarbeiter, Kooperationspartner etc.)

Schutz nach dem Geschäftsgeheimnisgesetz

1. Schutzanforderungen nach § 2 Nr. 1 Geschäftsgeheimnisgesetz
   (1) Informationen sind geheim
   (2) Information haben wirtschaftlichen Wert
   (3) Angemessene Schutzmaßnahmen

2. Ausschluss durch das Kriterium der Geheimhaltung
   - Strukturierte Datensets und derivative Daten können auch bei bekannten Rohdaten (z. B. Topographische Daten, Verkehrsdaten) geheim sein

3. Ausschluss durch das Kriterium des wirtschaftlichen Wertes?
   - Fehlender wirtschaftlicher Wert von kontextlosen Einzeldaten (Temperaturwerte etc.)
   - Potentieller wirtschaftlicher Wert bei Datensets, soweit nicht belanglos

Schutz nach dem Geschäftsgeheimnisgesetz

4. Ausschluss durch das Kriterium der angemessenen Maßnahmen?

   4.1 Wegfall des Schutzes bei Cloud-Lagerung der Daten oder Outsourcing? (-)

   4.2 Wegfall des Schutzes der Daten durch Zugänglichmachung in Entwicklungskooperationen?
   - (-) bei Geheimhaltungsverpflichtungen
   - (+) bei späteren Offenbarungen in der weiteren Wertschöpfungskette der Daten

   4.3 Maßnahmen der IT Sicherheit sind häufig auch als Indiz für angemessene Geheimhaltungsmaßnahmen zu betrachten

Vertragliche Geheimhaltungsregelungen

1. Häufig Grundlage von differenzierten Datennutzungsmodellen

2. Gestaltung vertraglicher Geheimhaltungsmodelle
   - 2.1 *Keine* Fiktion der Geheimhaltung
   - 2.2 Vertraglicher Gestaltungsspielraum für erlaubte und unerlaubte Nutzungen (Datentransferierungen und Nutzungen)

3. Durchbrechung von vertraglichen Geheimhaltungspflichten
   - Veröffentlichungspflichten im Bereich der staatlich geförderten Forschung
   - Regulatorische Transparenzpflichten (Produktsicherheit etc.)

Probleme des Schutzes von Daten durch Geheimhaltung?

1. Schutz setzt einen dauerhaften „Aggregatzustand“ der Geheimhaltung voraus

   - Unterschied zum Schutz von KI-Verfahren durch Patentschutz oder Softwareentwicklungen nach § 69a UrhG

2. Vertragliche Geheimhaltungspflichten binden nur Vertragspartner

   - Verletzung des Systems führt (nur) zu Schadensersatzanspruch gegenüber dem Vertragspartner

3. Risikoerhöhung vertragsbasierter Datenschutzmodelle *mit jeder Ausweitung des Kreises der Datenzugangsberechtigten*

Grenzen des Geheimhaltungsschutzes als Grundlage von Datenmodellen

1. Fehlende Eignung des Geheimhaltungsschutzes für datenbasierten Geschäftsmodells, die zwingend auf einer Offenbarung beruhen

2. Weitere Risiken durch Schutzkonzepte, die auf einer Geheimhaltung beruhen

   2.1 Monopolisierungsrisiken durch Akkumulierung von geheimen Daten (Spezifische Daten in der Pharmaforschung etc.)

   2.2 Transparenzprobleme (Versicherungsmodelle, Smart Energy etc.)

   2.3 Innovationsblockierung

Sui generis Schutz für Daten?

1. Diskussion zu einem neuen Leistungsschutzrecht für Daten
   - Nachteilige *Geheimhaltungsschutz ist aktuell wesentliche rechtliche Schutzform für Daten, welche nicht die urheberrechtliche Schöpfungshöhe überschreiten bzw. in Datenbank aufgehen*

2. Anknüpfungspunkte für die Gewährung eines rechtlichen Schutzes?

   2.1 Technische Entwicklungsleistungen oder Schöpfungshöhe? (-)

   2.2 Aufwand für die Erhebung und gegebenenfalls Aufbereitung der Daten

Sui generis Schutz für Daten?

3. Reichweite eines möglichen neuen Schutzes?

   3.1 Ausschließlichkeitsrechte würden die Datennutzung blockieren

   3.2 Nur Anspruch auf Gegenleistungen

   - Finanziell oder in Gewährung eigener Rechte

3. Aktuelle Verordnungsinitiativen auf der EU Ebene spricht eher gegen die Einführung eines eigenen Rechtes (= Sonderregelungen für Datennutzung, jedoch keine Sonderregelung für den Schutz)

III. Rechtezuordnung und Nutzungsrechtsregelungen in Entwicklungskooperationen

Rechtezuordnung in Bezug auf die Daten

1. Befund: Hauptinstrument der Zuordnung der Inhaberschaft und der Verwertungsrechte ergibt sich aus Verträgen, da gesetzliche Regelungen nicht hinreichend sind

2. Konfliktpotential verlangt nach vertraglicher Regelungen
   - 2.1 Berücksichtigung der unterschiedlichen Datenebenen (Rohdaten, Trainingsdaten, derivative Daten)
   - 2.2 Berücksichtigung der unterschiedlichen rechtlichen Schutzform (Geheimhaltung, Datenbankwerke etc.)
   - 2.3 Haftungsfragen für Datenverwendung

Dateninhaberschaft in Entwicklungs- und Kooperationsprojekten

1. Notwendige Differenzierung der Regelungsgegenstände

   1.1 Background

   - Rechte an *eingebrachten* Rohdaten
   - Rechte an *eingebrachten* Trainingsdaten

   1.2 Foreground

   - Rechte an den im Projekt neu generierten Rohdaten (Versuchs- oder Testdaten im Entwicklungsprojekt)
   - Rechte an im Projekt neu trainierten oder weiter trainierten KI-Trainingsdaten

Dateninhaberschaft in Entwicklungs- und Kooperationsprojekten

2. Differenzierung der rechtlichen Regelungsform

   2.1 Regelung der *Datenzuordnung*

   2.2 Regelung der *Datennutzung*

Regelungsprinzipien bei eingebrachten Roh-daten oder Trainingsdaten

1. Keine Inhaberschafts*übertragung*, soweit dies nicht ausdrücklich intendiert ist

2. Bestimmung der Nutzungsrechte
   - 2.1 Nutzungsrechte für Projektrealisierung
   - 2.2 Weitergehende Nutzungsrechte für die Verwertung?
     - (1) Nutzungen im Zusammenhang mit dem gemeinsam entwickelten KI-Systemen
     - (2) Nutzung für andere KI-Systeme oder sonstige Zwecke?

Regelung zu im Projekt neu generierten Roh-daten und Trainingsdaten

1. Allein- oder Mitinhaberschaft?
   - Probleme der „Flucht“ in die Mitinhaberschaft

2. Nutzungsrechte

2.1 Nutzung für *Projektrealisierung*

2.2 Nutzung für *gemeinsame* Verwertungsformen

2.3 Nutzung für sonstige Verwertungen

3. Kompensationsregelungen zur Steuerung des Interessenausgleichs

„Überlagerung“ der Regelungen zu den Daten durch andere Rechte zum Datenhandling?

1. Regelung zu den Datennutzungsmodellen hat auch die Rechtslage hinsichtlich der datenverwendenden Systeme zu berücksichtigen
   - Patentierungswettlauf hinsichtlich der *KI-Systeme, KI-Methoden und Verwendung der KI-generierten Ergebnisse*

2. Interoperabilität hinsichtlich der Datenformate, der Datenübertragungsformen, der Datenverarbeitung und datenbasierten Steuerung?

3. Grenzen der Standardisierung aufgrund proprietärer Geschäftsmodelle
   - Aktuelle Dominanz proprietärer Modelle in KI-Anwendungsgebieten (Smart Home etc.)

Rechtliche „Infizierung“ durch den Daten-austausch in KI-Entwicklungskooperationen

1. Integrierung von „infizierten“ Daten anderer Partner kann rechtliche Risiken erhöhen

   1.1 Verstoß gegen Geheimhaltungsverpflichtungen, Exklusivitätsregelungen oder sonstige Rechte Dritter hinsichtlich der von dem Kooperationspartner bereitgestellten Daten

   1.2 Verstoß gegen Datenschutzbestimmungen etc.

2. Freistellungsregelungen schaffen keine hinreichende Abschirmung gegenüber Ansprüchen Dritter

   – Kontrollfrage: Verfügt der Kooperationspartner über hinreichende rechtliche Sensibilität bei der Datenbeschaffung?

IV. Haftung für Daten und Regulierung der Datennutzung

Verletzung von Rechten Dritter bei der Nutzung von Daten?

1. Soweit kein rechtlicher Schutz der Daten besteht, ist eine Nutzung von Daten möglich
   – Grundsatz der Freiheit von Fakten und Informationen

2. Besonderer rechtlicher Schutz?
   (1) Vertragliche Bindungen; Copyleft/share-alike Verpflichtungen
   (2) Schutzrechte / Wettbewerbsrecht
   (3) Geschäftsgeheimnisgesetz
   (4) Urheberrechtlicher Schutz / Datenbankschutz
   (5) Persönlichkeitsrechte nach § 823 BGB
   (6) Recht am eigenen Bild nach dem Kunsturhebergesetz (KUG) und § 201a StGB
   (7) Spezieller Unterlagenschutz
   (8) Datenschutz

Urheberrechtsverletzung beim Data Mining und bei der Gewinnung von Trainingsdaten

1. Identifizierung und Extrahierung von Daten sowie Transformierung in ein maschinenlesbares Datenformat

2. Soweit urheberrechtlich geschützte Texte, Bilder oder sonstige Werke betroffen sind, dürfte zumindest das Vervielfältigungsrecht nach § 16 UrhG betroffen sein, soweit nicht ein reines „crawlen“ erfolgt

3. Schrankenregelung für das Data Mining
   - 3.1 zu wissenschaftlichen Zwecken nach § 60d UrHG
   - 3.2 für kommerzielle und sonstige Zwecke nach § 44b UrhG
     - (1) Löschungspflicht nach Wegfall der Erforderlichkeit nach § 44b Abs. 2 UrhG
     - (2) „Opt-out“ Möglichkeit nach § 44b Abs. 3 UrhG (maschinenlesbar zumindest bei online zugänglichen Werken)

Verstoß gegen das Datenschutzrecht bei der Nutzung von personenbezogenen Daten

1. Alle Informationen, „die sich auf eine identifizierte oder identifizierbare natürliche Person“ beziehen (Art. 4 Nr. 1 DS-GVO)

2. Was ist „identifizierbar“?
   - Es kommt nicht allein auf die Intention hinsichtlich der Identifizierung an, sondern ob eine Identifizierung nach allgemeinem Ermessen *wahrscheinlich* ist
   - = (praktische und rechtliche Begrenzungen; Möglichkeit der Heranziehung von Zusatzwissen etc.)

Verstoß gegen das Datenschutzrecht bei der Nutzung von personenbezogenen Daten

1. Daten müssen auf rechtmäßige Weise, nach Treu und Glauben und in nachvollziehbarer Weise verarbeitet werden (Art. 5 Abs. 1 lit. a) DSGVO)
   - Insbesondere Grundsatz der Transparenz

2. Prinzip der Zweckbindung (Art. 5 Abs. 1 lit. b) DS-GVO)
   - Personenbezogene Daten dürfen nur für festgelegte, eindeutige und legitime Zwecke erhoben und weiterverarbeitet werden

3. Prinzip der Datenminimierung (Art. 5 Abs. 1 lit. c) DS-GVO)

Verstoß gegen das Datenschutzrecht bei der Nutzung von personenbezogenen Daten

4. Prinzip der Speicherbegrenzung (Art. 5 Abs. 1 lit. e) DSGVO)
   - Höchstdauer der Speicherung wird durch die Zweckerfüllung begrenzt

5. Prinzip der Integrität und Vertraulichkeit (Art. 5 Abs. 1 lit. f) DSGVO)
   - Gewährleistung der Sicherheit der Daten

6. Rechenschaftspflicht nach Art. 5 Abs. 2 DSGVO
   - Nachweispflicht

7. Prinzip der Einwilligung nach Art. 6 Abs. 1 lit a. DSGVO

Datenschutzrechtliche Sonderprobleme

1. Datengewinnung aus verschiedenen Quellen (Big-Data-Analysen etc.)
   - Zusammenführung von Daten ist nach dem Datenschutzrecht *Datenneuerhebung (= Unterrichtungspflicht des Betroffenen)*
   - *Ausnahmen nach Art. 14 DS-GVO: - Informationen zur zweiten Erhebung vorhanden, gesetzlich geregelt, unmöglich oder unverhältnismäßig?*

2. Spezifische Anwendungsbereiche (z.B. Gesundheitsdaten für wearables oder health Projekte)
   - Gesundheitsdaten Art. 4 Nr. 15 DS-GVO = Besondere Anforderungen
   - Getrennte bzw. gemeinsame Verantwortung bzw. Auftragsverarbeitung z.B. bei Ausstattung eines wearables mit App/Software?
   - Hohe Anforderungen an Einwilligungen (Widerruflichkeit!)

Datenschutzrechtliche Sonderprobleme

3. Privilegierung der wissenschaftliche Forschung nach Art. 89 DS-GVO
   - 3.1 Wissenschaftliche Forschung ?
   - 3.2 Öffentliches Interesse?
   - 3.3 Erforderlichkeit?
   - 3.4 Datenminimierung / Anonymisierung bzw. Pseudonymisierung

4. Anonymisierung - Lösung aller Probleme?
   - 4.1 Für das konkrete Projekt machbar?
   - 4.2 Re-Identifizierbarkeit bei Big Data Sachverhalten durch hohe Kontextdichte

Weitergehende Haftung für Rohdaten und Trainingsdaten

1. Aktuell keine offizielle gesetzliche Regelung
   - Entwurf der EU-Kommission vom 21. April 2021 einer Verordnung zur Regulierung der Nutzung Künstlicher Intelligenz (Artificial Intelligence Act) sieht in Art. 10 Vorgaben für Trainingsdaten bei „Hochrisiko-KI-Systemen“ vor
   - Ansatz der Regulierung durch Begründung von Haftungsrisiken

2. Vertragliche Haftung zwischen Datenlieferanten und Geschädigten
   - 2.1 Kaufrechtliche Haftung nach § 433 ff. BGB
   - 2.2 Anwendung der schenkungsrechtlichen Haftungsreduktion bei unentgeltlicher Datenüberlassung

Haftung für Rohdaten und Trainingsdaten

2.3 Zumeist keine Vertragsbeziehung zwischen Datenlieferanten/Datenveredlern und Geschädigten (z. B. Nutzern autonomer Systeme, Anwendern von medizinischen Systemen etc.)

3. Anwendbarkeit des Produkthaftungsgesetzes?

3.1 Anwendbarkeit auf trainiertes KI-System und Software (+)

3.2 Anwendbarkeit auf Daten: strittig

Haftung für Rohdaten und Trainingsdaten

4. Deliktische Haftung nach § 823 BGB
   - 4.1 Verkehrssicherungspflicht bei dem Inverkehrbringen von Trainingsdaten für Hochrisiko-KI-Systeme (autonome Systeme etc.)?
   - 4.2 Schwierige Beweisführung bei komplexen Kausalketten

5. Regelung der Verantwortung in KI Entwicklungsgemeinschaften und Freistellungsregelungen
   - 5.1 Umfassende Regelung zur Verantwortung im Entwicklungsprozess
     - Datenzusammenführung, Aufbereitung
   - 5.2 Umfassende Regelung zur Verantwortung hinsichtlich der jeweiligen oder gemeinsamen Datennutzung

V. „Open“- Modelle und Anspruch auf Datennutzung?

Regelung von Datennutzungsmodellen - Zwischen proprietären Systemen und „Open“-Modellen

1. „Open“ setzt zunächst Schutz voraus – Keine Rechteanmaßung!

2. „Open“ Lizenzen
   - 2.1 Open Source (Software), z.B. General Public License (GPL)
   - 2.2 Open Content (Texte, Bilder oder sonstige Inhalte), z.B. Creative Commons (CC)
   - 2.3 Open Data (Datenbanken), z.B. Open Data Base License, ODbL
   - 2.4 Public Domain Erklärungen

Regelung von Datennutzungsmodellen

3. Analyse der Rechteinhalte
   - Non Commercial (NC), share-alike bzw. Copyleft etc.

4. Nachteile in Datennutzungsmodellen
   - 4.1 Verlust von Anreizen für Datenbereitsteller und Datenveredler?
   - 4.2 Kollision mit proprietären Geschäftsmodellen und Regelungskonflikte bei der Einbindung in solche Geschäftsmodelle?
   - 4.3 Kollision bei der Kombination von „Open“ Daten und proprietären Daten und KI-Trainingsmodellen etc.

Open - Modelle

1. Open - Plattformen im Bereich des maschinellen Lernens
   - Es entwickeln sich KI-spezifische Lizenzmodelle, welche wertebasiert sind (siehe z. B. Responsible AI Licenses unter www.licenses.ai)
   - Anbieter von Daten z.B. Kaggle, Google Data Set Search, Deep Drive BDD 100 K setzen auf „Open“ Modelle

2. Viele Unternehmen, die in proprietären Systemen denken, stellen KI-Tools unter Open Source-Bedingungen bereit, um KI-Entwicklungen zu stimulieren
   - z. B. TensorFlow von Google unter einer Apache-2.0 Open Source-Lizenz; PyTorch von einem Facebook-Forschungsteam, welches auf Grundlage einer Open Source-Lizenz angeboten wird

Anspruch auf Gewährung eines Datenzugangs

1. Mögliche rechtliche Grundlagen

   1.1 Vertragliche Ansprüche

   1.2 Kartellrechtlicher Anspruch auf Gewährung eines Datenzugangs, insbesondere bei einer marktbeherrschenden Stelle des Dateninhabers

   1.3 Wettbewerbsrecht (UWG)

   - Spezifizierung eines „unlauteren Verhaltens“ im Falle der Verweigerung eines Datenzugangs

   1.4 Bestehende und zukünftige rechtliche Neuregelungen zur Sicherung eines Datenzugangs

Anspruch auf Gewährung eines Datenzugangs

2. Differenzierung hinsichtlich der Anspruchsinhalte

   2.1 Anspruch auf Datennutzung

   2.2 Abgrenzung gegenüber dem Anspruch auf Zugang zur Sicherung einer Transparenz und Kontrolle

3. Nachteilige Wirkung eines Anspruchs auf Zugang und Nutzung

   3.1 Starke Beeinträchtigung von Anreizwirkungen bereits durch das Risiko eines möglichen Anspruchs

   3.2 Fehlende Heilbarkeit von Fehlentscheidungen

Neue Erweiterungen des Datenzugangs?

1. Open Data-Richtlinie (RL GU 2019/1024)

   - Öffentliche Stellen sowie öffentliche Unternehmen Informationen müssen Daten des öffentlichen Sektors (PSI-Daten (Public Sector Information)), insbesondere für den KI-Datenanalysebereich (z.B. Klimaforschung, Mobilität etc.) bereitstellen

2. *Aktionsplan Forschungsdaten des BMBF/BMWK und der EU*

   - Nationale Forschungsdateninfrastruktur (NFDI), Cloud- und Dateninfrastruktur (GAIA-X)-Förderung(?), European Open Science Cloud (EOSC)

3. EU Digital Markets Act (24.3.2022) und GWB-Digitalisierungsgesetz (18.1.2021)

   Unternehmen, die z. B. als „Gatekeeper“ im Zugang zu bestimmten Daten agieren, unterliegen strengeren Regeln.

Auswirkungen des EU Data Act (Kommissions-entwurf vom 23.2.2022)

1. Recht der Nutzer von vernetzen Geräten auf Zugang zu den von ihnen erzeugten Daten sowie auf Nutzung solcher nutzergenerierten Daten (gilt B2B sowie B2C)

2. Etablierung eines FRAND-Standards für Datenlizenzverträge (faire Datenzugangsregelung)

3. Recht auf Datenzugang durch öffentliche Stellen, soweit dies im öffentlichen Interesse geboten ist

4. Verpflichtung zur Umsetzung von offenen Standards und Schnittstellen zur Erleichterung eines Anbieterwechsels

Regulierung durch den EU Data Act

5. Auswirkungen für Entwicklungsgemeinschaften:

   5.1 Der Data Act betrifft industrielle Daten und nicht personenbezogene Daten, jedoch werden Regelungsansätze aus dem Datenschutzrecht genutzt (weitgehende Entscheidungsmacht des Datenbereitstellers)

   5.2 Kein Eigentum oder neue Schutzform für Daten, sondern Fokussierung auf Datenzugang

   5.3 Zementierung des Grundsatzes eines fairen Datenzugangs

Vielen Dank!