Deutsche Bank
Chief Administrative Office

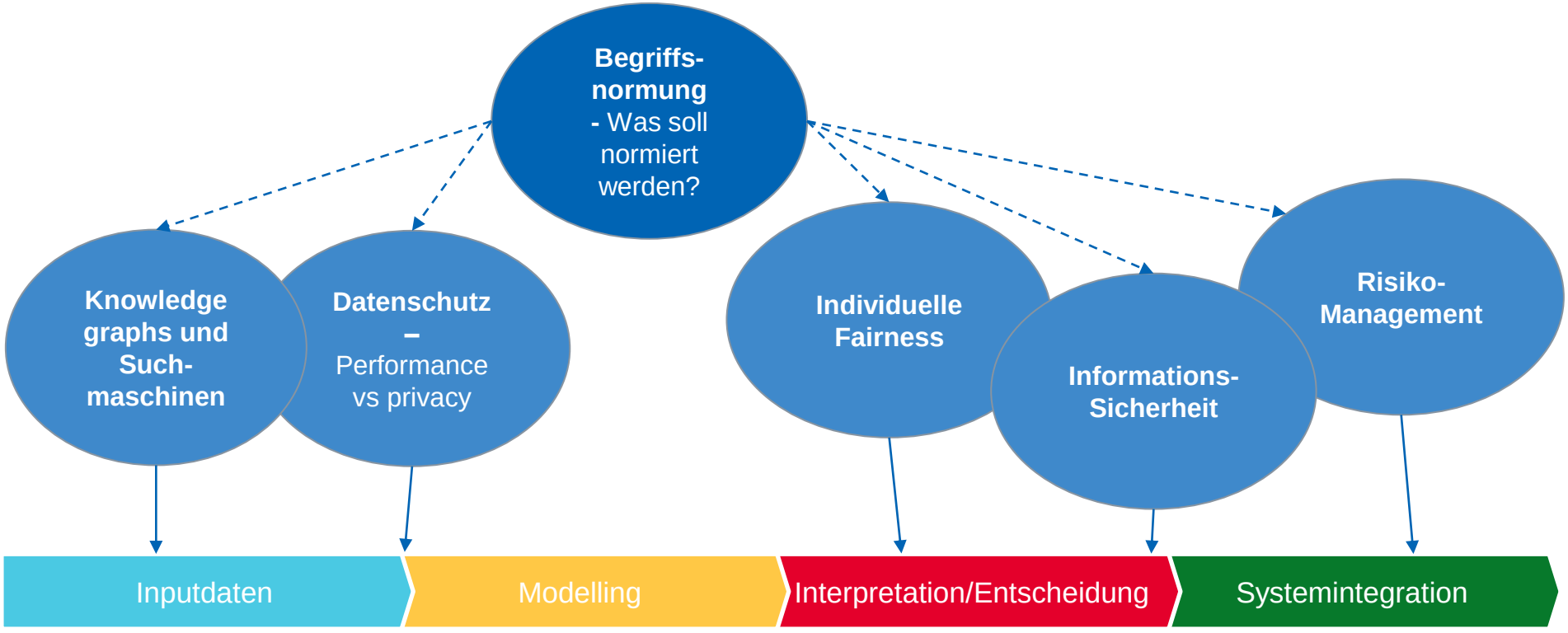
KI und Normung im Finanzsektor

DIN Workshop

Adrian Seeliger, Oliver Maspfuhl, AG Finanzdienstleistungen,
Frankfurt, 18. April 2023

KI-Regulierung – diesmal richtig
Welche Aspekte sind finanzspezifisch?

Normungsroadmap KI II – AG Finanzdienstleistungen



KI-Regulierung – diesmal richtig
Welche Normungsbedarfe wurden identifiziert?

Normungsbedarfe Finanzdienstleistungen

Code	Bedarf
08-01	Definition nachprüfbarer Antidiskriminierungsmetriken zum Nachweis der Diskriminierungsfreiheit einer KI-Lösung
08-02	Normung der für Nichtdiskriminierung relevanten Merkmale und des Umgangs damit
08-03	Normung der Berücksichtigung von Nichtdiskriminierungsaspekten bei der Erstellung einer KI-Lösung zum Nachweis der Diskriminierungsfreiheit
08-04	Definition des Begriffs Fairness durch nachprüfbare Metriken
08-05	Regeln für den Nachweis der Abdeckung aller relevanter Faktoren bei Gruppenbetrachtungen
08-06	Erarbeitung und Definition von (Mindest-)Anforderungen an eine KI-Plattform
08-07	Rahmenbedingungen zum Umgang mit Trainingsdaten für KI-Modelle
08-08	KI-spezifische Angriffsszenarien und Schutzmaßnahmen
08-09	Festlegung von Kriterien, die für eine automatisches Entity-Matching ausreichend sind
08-10	Festlegung von Kriterien, wie die Verlässlichkeit von Matching mithilfe von statischen Modellen gemessen werden kann und welche Mindestwerte notwendig sind.
08-11	Festlegung von Mechanismen, mit denen die Nutzer*innen die Verwendung der eigenen Identität überwachen können
08-12	Leitfaden Usable Security
08-13	Vorgehensweise für die Sicherheitsbetrachtung relevanter Stakeholder
08-14	Normen für die Validierung des Modells, um bewerten zu können, ob das KI-System für den Einsatz in der produktiven Umgebung hinreichend überprüft wurde UND Regeln für die regelmäßige Re-Evaluierung von KI-Systemen aufstellen
08-15	Normen für die Transparenz zur Fehlerkorrelation des Systems
08-16	Definition hinreichender Maße für Transparenz, damit der Entwickler weiß, welche zusätzlichen Informationen bereitgestellt werden müssen, um die entsprechende Architektur des KI-Systems zu konstruieren
08-17	Normung von Dokumentationspflichten zum Ursprungskontext von Modellen und (Trainings-)Daten
08-18	Normen für die Transparenz zur Konfidenz und Modellrisiken von Einzelentscheidungen

KI-Regulierung – diesmal richtig
Welche Normungsbedarfe sind auch außerhalb derFinanzwelt wichtig?

Normungsbedarfe Finanzdienstleistungen – ausgewählte spezifische Beispiele

Code	Bedarf	Inhalt und Bedarfserläuterung
08-02	Normung der für Nichtdiskrimi-nierung relevanten Merkmale und des Umgangs damit	In den Gesetzen und Vorgaben zu Antidiskriminierung werden die relevanten Merkmale inkonsistent genannt. Beispiele: Charta der Grundrechte der EU (Art. 21 „Nichtdiskriminierung“): „Diskriminierungen, insbesondere wegen des Geschlechts, der Rasse, der Hautfarbe, der ethnischen oder sozialen Herkunft, der genetischen Merkmale, der Sprache, der Religion oder der Weltanschauung, der politischen oder sonstigen Anschauung, der Zugehörigkeit zu einer nationalen Minderheit, des Vermögens, der Geburt, einer Behinderung, des Alters oder der sexuellen Ausrichtung, sind verboten.“ Vertrag über die Arbeitsweise der EU: „... discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.“ Eine einheitliche und abschließende Liste der Merkmale kann helfen, Aufwände bei der Erstellung von KI-Lösungen zu vermeiden bzw. die Leistungsfähigkeit einer KI-Lösung zu verbessern. Darüber hinaus soll genormt werden, wie die relevanten Merkmale bei der Erstellung der KI-Lösungen berücksichtigt werden sollen. Ein genereller Ausschluss ist möglicherweise kontraproduktiv. Beispiel: Unter der Annahme, dass die Kreditwürdigkeit einer Person von der Dauer der bisherigen Bankverbindungen abhängt und zugleich historisch bedingt insbesondere ältere Frauen im Mittel kürzere Bankverbindungen haben, wäre eine gegebene Dauer einer Bankverbindung für eine Frau möglicherweise positiver zu werten als für einen Mann. Entfernt man das Merkmal „Geschlecht“ aus den Lerndaten, wären ältere Frauen systematisch benachteiligt. Anbieter*innen und Entwickler*innen von KI-Lösungen profitieren von der Rechtssicherheit durch konsistente Regeln.
08-05	Regeln für den Nachweis der Abdeckung aller relevanter Fakto-ren bei Gruppenbetrach-tungen	Wenn KI-Systeme Aussagen über Gruppen machen, sind diese nicht notwendigerweise auf das Individuum übertragbar. Daher muss sichergestellt sein, dass entweder keine wesentlichen individuellen Faktoren im Modell fehlen oder eine Geltendmachung und Berücksichtigung grundsätzlich möglich ist, sofern sie nicht ethischen Grundsätzen widerspricht. Dies gilt insbesondere, wenn Grundrechte aufgrund von Modellen eingeschränkt werden, die Aussagen über Gruppen von Individuen machen. Im Kontext von Finanzanwendungen, aber auch bei anderen sozioökonomischen Systemen, steht häufig eine Risikobetrachtung über die Gruppe im Vordergrund, etwa bei der Vorhersage des erwarteten Verlusts in einem Kreditportfolio oder bei der erwarteten Ausbreitung einer Krankheit. Eine korrekte Vorhersage für das Portfolio und entsprechende Risikopreise (oder, analog, entsprechende Gesundheitsschutzmaßnahmen), muss aber auch für das Individuum (dessen Grundrechte berührt werden) unter allen für es verfügbaren Informationen optimiert werden. Das heißt, es müssen je nach Schwere der Konsequenzen alle individuellen Faktoren berücksichtigt werden, die nachweislich einen signifikanten Einfluss auf die Prognose haben. Es braucht daher Regeln, nach denen die relevanten Faktoren bestimmt werden.

KI-Regulierung – diesmal richtig
Welche Normungsbedarfe sind auch außerhalb derFinanzwelt wichtig?

Normungsbedarfe Finanzdienstleistungen – ausgewählte spezifische Beispiele

Code	Bedarf	Inhalt und Bedarfserläuterung
08-09	Festlegung von Kriterien, die für eine automatisches Entity-Matching ausreichend sind	Für kritische Systeme dürfen Identitäten in zwei unterschiedlichen Datensätzen nur gematcht werden, wenn sie zu 100 % übereinstimmen. Daher muss festgelegt werden, welche Kriterien hierfür ausreichend sind. Auch für nicht-kritische Systeme dient es der Qualität, wenn Daten den richtigen Identitäten zugeordnet werden. Beispiel: Kundennummer ist nicht eindeutig zur Person zuzuordnen. Im Finanzsektor sind die Datensätze, die zum Training einer KI verwendet werden, nicht immer über eindeutige Identifizierungsmerkmale zugeordnet wie z. B. die Personalausweisnummer oder die Krankenversichertennummer im Gesundheitssektor.
08-10	Festlegung von Kriterien, wie die Verlässlichkeit von Matching mithilfe von statischen Modellen gemessen werden kann und welche Mindestwerte notwendig sind.	Wenn Identitäten nur probabilistisch gematcht werden, muss gemessen werden können, wie verlässlich das Matching ist und für welche Art der Anwendung welche Mindestverlässlichkeiten gelten sollen. Die falsche Zuordnung von Daten zu Entitäten ist ebenso eine Fehlerquelle für Training und Anwendung von KI wie die Fehlerhaftigkeit von korrekt zugeordneten Daten.
08-15	Normen für die Transparenz zur Fehlerkorrelation des Systems	Ein KI-System soll in standardisierter Weise transparent machen, wie die Korrelationsstruktur der statistischen Unsicherheiten aussieht. Statistische Unsicherheiten der Ausgaben eines KI-Systems sind nicht notwendigerweise unabhängig. Für das Risikomanagement möglicher Fehler des Systems ist eine Kenntnis der Abhängigkeitsstruktur entscheidend. Zudem muss definiert werden, inwiefern ein Input unter Unsicherheit erstellt wurde (durch ein vorgeschaltetes Modell oder einen Datensatz).