

## ANFORDERUNGEN AN EIN QM-SYSTEM NACH KI-VO AM BEISPIEL EINER KI-BASIERTEN BEWERBERAUSWAHL

**Dr. Fabian Küppers, Thorsten Gedicke,  
RAin und SyndikusRAin Johanna Schneider**  
e:fs TechHub GmbH Gaimersheim

**RA Christian Galetzka,  
LL.M.**  
Jun Rechtsanwälte Würzburg

## Einleitung

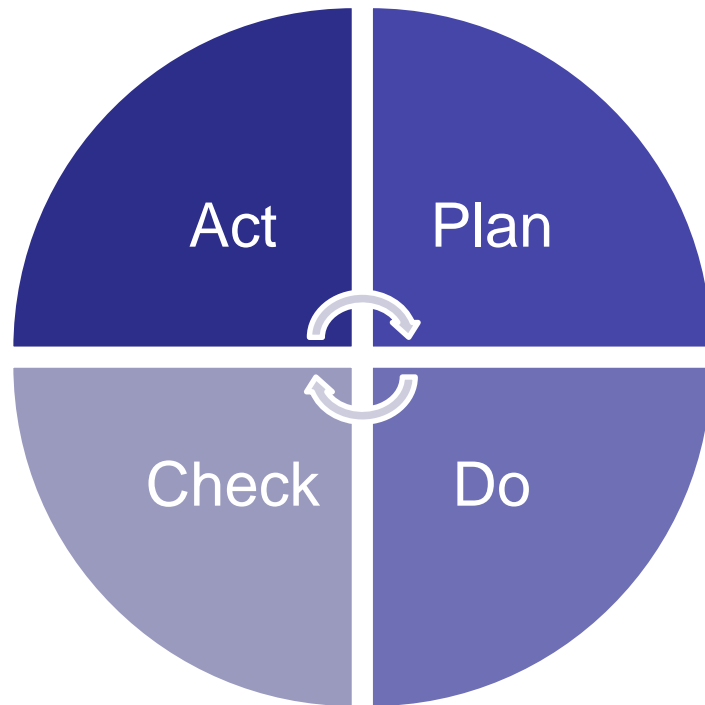
Gemäß ISO 9000:2015 wird das Qualitätsmanagementsystem (kurz: QMS) wie folgt definiert:

„Ein QM-System umfasst Tätigkeiten, mit denen die Organisation ihre Ziele ermittelt und die Prozesse und Ressourcen bestimmt, die zum Erreichen der gewünschten Ergebnisse erforderlich sind. Das QMS führt und steuert in Wechselwirkung stehende Prozesse und Ressourcen, die erforderlich sind, um Wert zu schaffen und die Ergebnisse für relevante interessierte Parteien zu verwirklichen.“

Anforderungen an ein  
QM-System gem. Art.  
17 KI-VO



## QM-System



- ▶ Dokumentation in der Do-Phase zu erstellen
  - ▶ Einheitliche Definition „Fairness“
  - ▶ Robustheit: Betriebssicherheit, Informationssicherheit, Sicherheit des Systems vor gezielten Angriffen
  - ▶ Leistungsmetriken und Grenzen des Systems
  - ▶ Qualität der Datenbasis

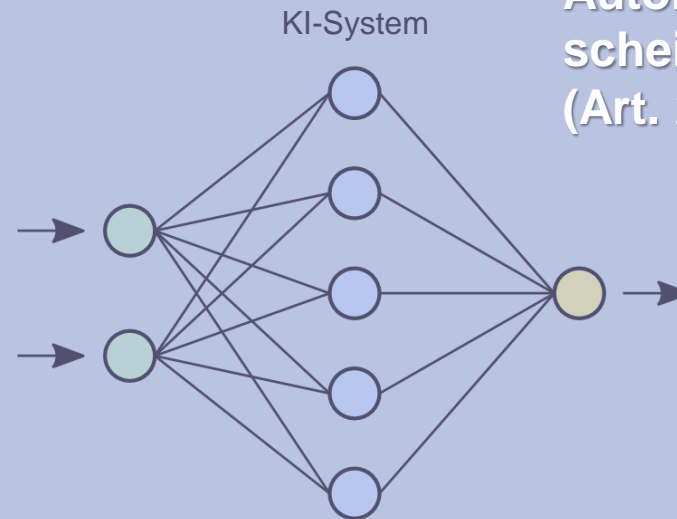
# Anwendungsbeispiel: KI-gestützte Bewerberauswahl

## KI-Recht

### Datenschutz und Arbeitsrecht



- Erfahrung
- Bildungsabschluss
- Gehaltsvorstellung
- Wohnort
- Hobbies



Automatisierte Entscheidungsfindung  
(Art. 22 DS-GVO)?

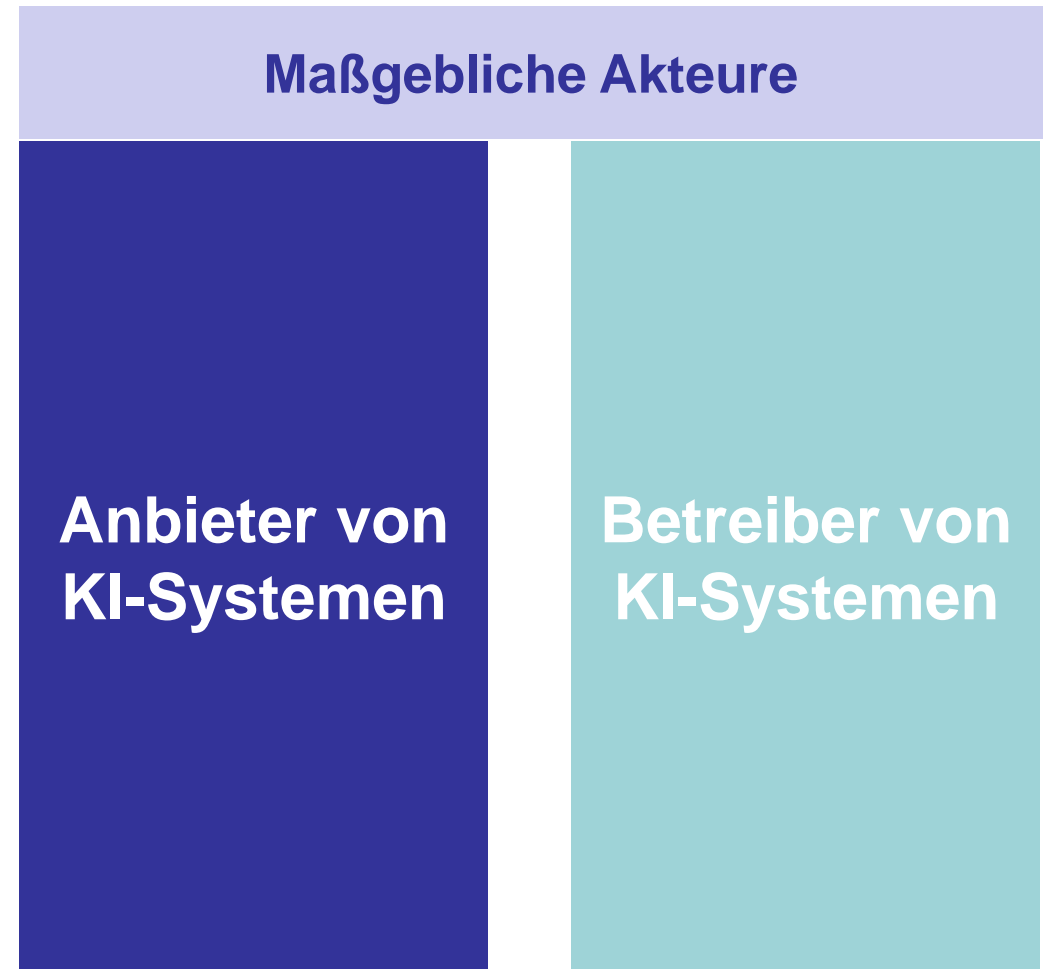
Einstellung ja/nein

Datenschutz, Arbeitsrecht, KI-Recht, Urheberrecht

# RECHTLICHE ANFORDERUNGEN AN KI-SYSTEME

# Rechtliche Anforderungen an KI-Systeme

- 1** **Datenschutz**  
Rechtmäßigkeit, Information, Risikomanagement, Dokumentation
- 2** **Arbeitsrecht**  
Rechtmäßigkeit, Diskriminierungsfreiheit, Information, Mitbestimmung
- 3** **KI-Recht**  
Rechtmäßigkeit, Fairness und Diskriminierungsfreiheit, Information, Risikomanagement, Dokumentation
- 4** **Urheberrecht**  
Schutz des Urheberrechts am Input und Output  
(eher bei generativen KI-System relevant)



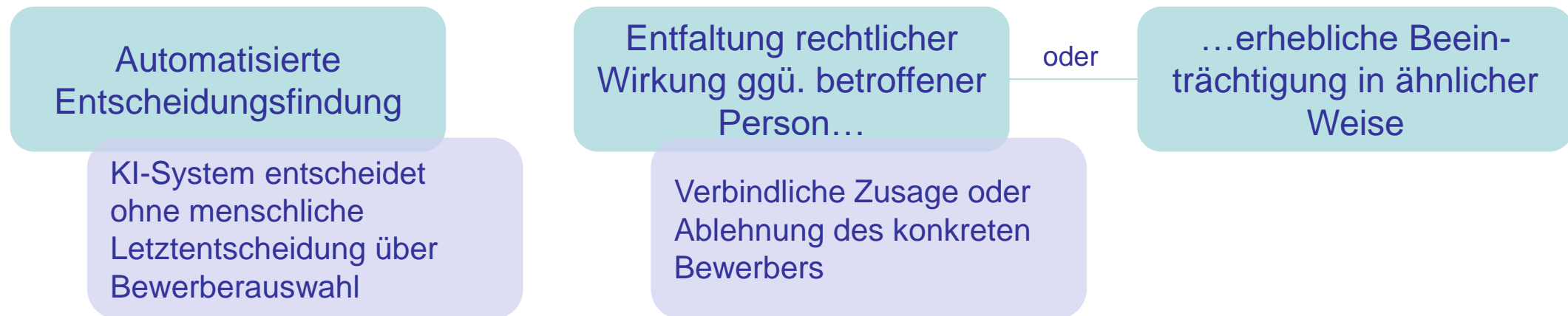
## Fokus auf einzelne datenschutzrechtliche Themenfelder

1 Automatisierte Entscheidungsfindung, Art. 22 DS-GVO

2 Informationspflichten nach DS-GVO und KI-VO

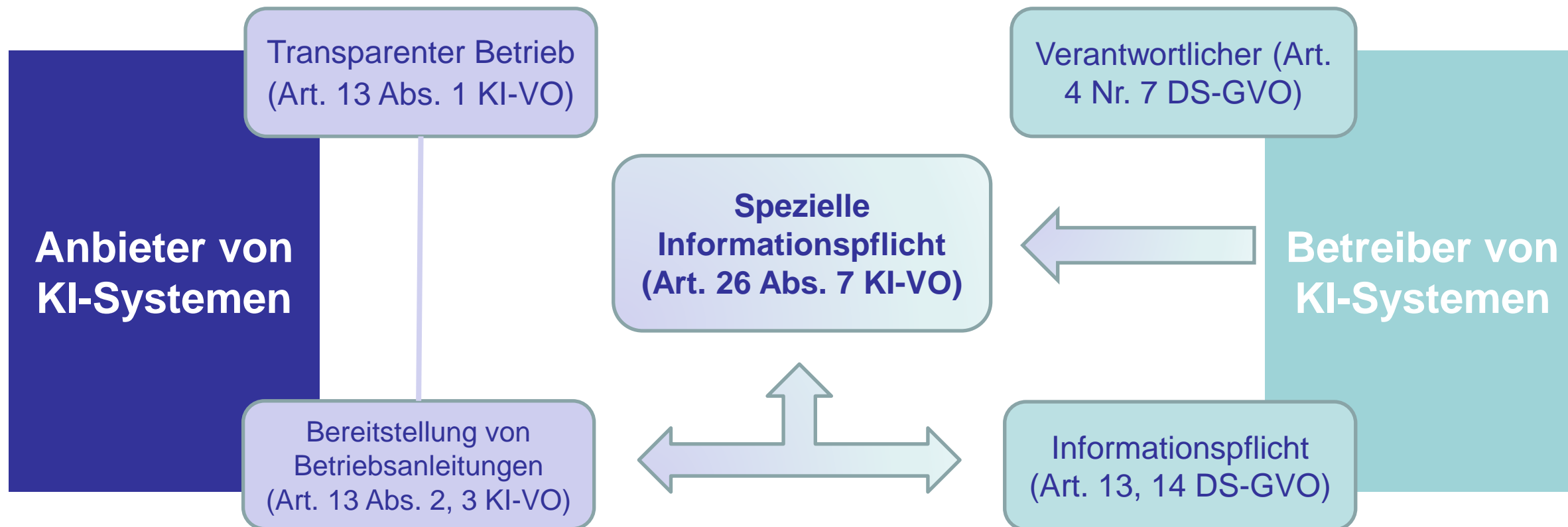
## Automatisierte Entscheidungsfindung, Art. 22 DS-GVO

- ▶ KI-gestützte personenbezogene Bewerberauswahl kann Profiling sein, vgl. ErwG 71 DS-GVO:  
„...Bewertung der persönlichen Aspekte einer Person [...], insbesondere zur Analyse oder **Prognose von Aspekten bezüglich Arbeitsleistung**...“
- ▶ Profiling fällt als **automatisierte Entscheidungsfindung** unter Art. 22 DS-GVO
- ▶ Grundsätzlich: **Einwilligung** der betroffenen Person erforderlich, auf Grundlage **transparenter Informationen**





## Informationspflichten nach KI-VO und DS-GVO



# TECHNISCHE ASPEKTE DER KI- BEWERTUNG

# Technische Anforderungen der KI-VO an Hochrisiko-KI

- 1 Technische Dokumentation**  
Dokumentation mit Nachweis der Konformität – Artikel 11 bzw. Anhang IV
- 2 Genauigkeit, Robustheit und Cybersicherheit**  
Bewertung anhand relevanter Metriken – Artikel 15
- 3 Transparenz und Information**  
Angabe der Leistungsfähigkeit und Leistungsgrenzen, auch in Bezug auf bestimmte Personengruppen – Artikel 13
- 4 Data Governance mit Bias-Untersuchung**  
Untersuchung und Bewertung der Daten auf Bias – Artikel 10
- 5 QM-System und Risikomanagement**  
Prozesse zur Erfüllung der Anforderungen sowie Risiko- und Systemmanagement – Artikel 9 und 17

## Aspekte der KI-Bewertung



### Datenqualität

Ermöglicht die Datengrundlage das zuverlässige Training und die Bewertung des KI-Modells?



### Performance

Wie leistungsfähig ist das KI-Modell?



### Robustheit & Konfidenz

Wie transparent und robust ist das KI-Modell bei Störungen und Unsicherheiten?

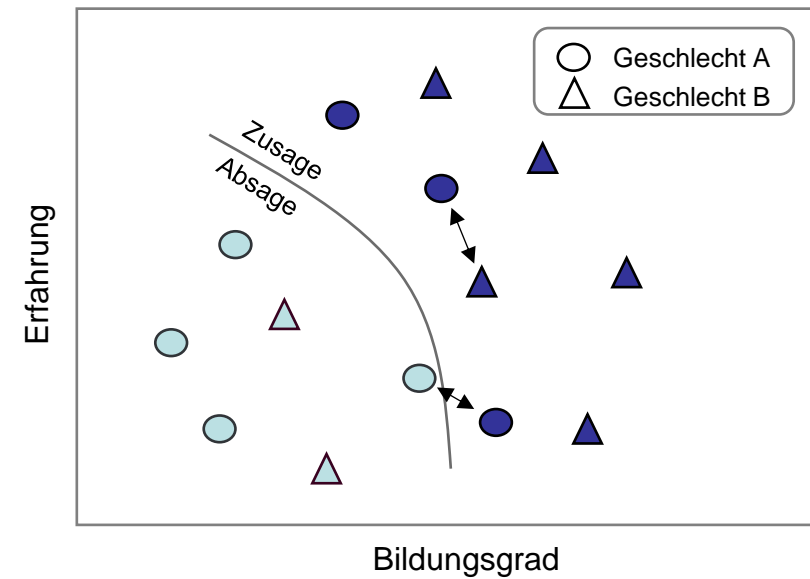
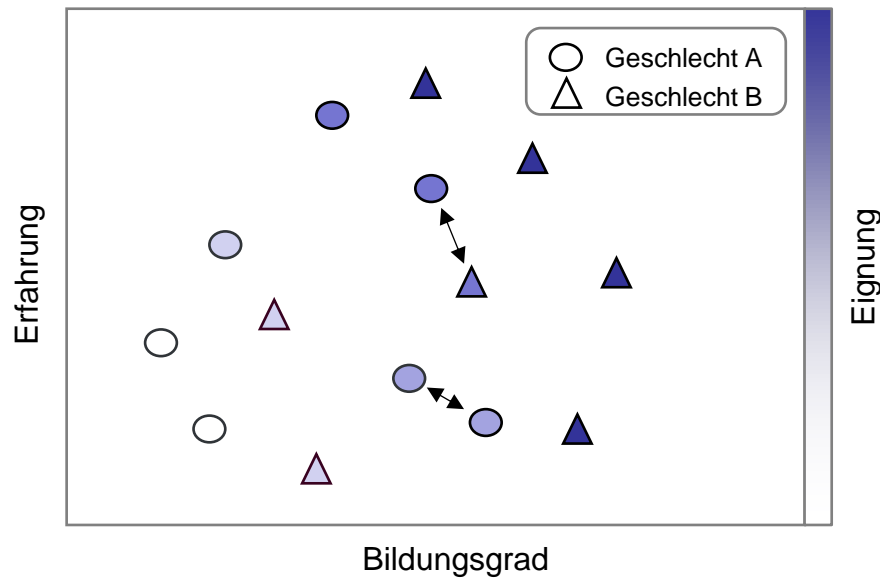


### Fairness

Ist das KI-Modell fair für alle Teile der Gesellschaft?

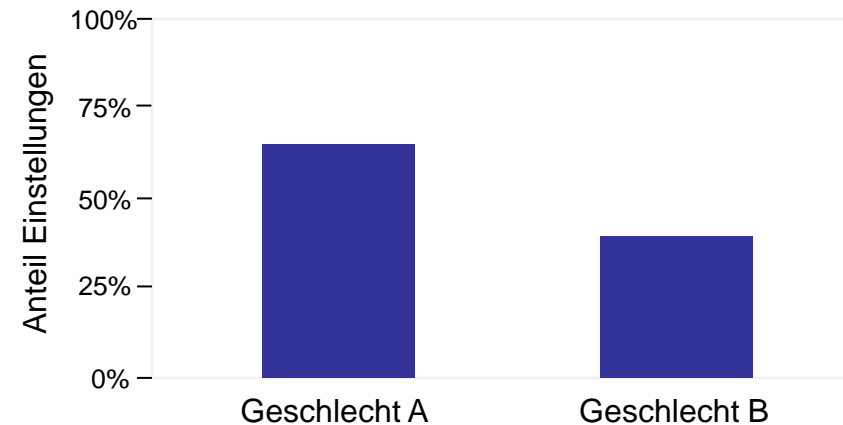
## Individual Fairness

- ▶ Ähnliche Bewertung von ähnlichen Individuen
- ▶ Übliches Ziel im maschinellen Lernen



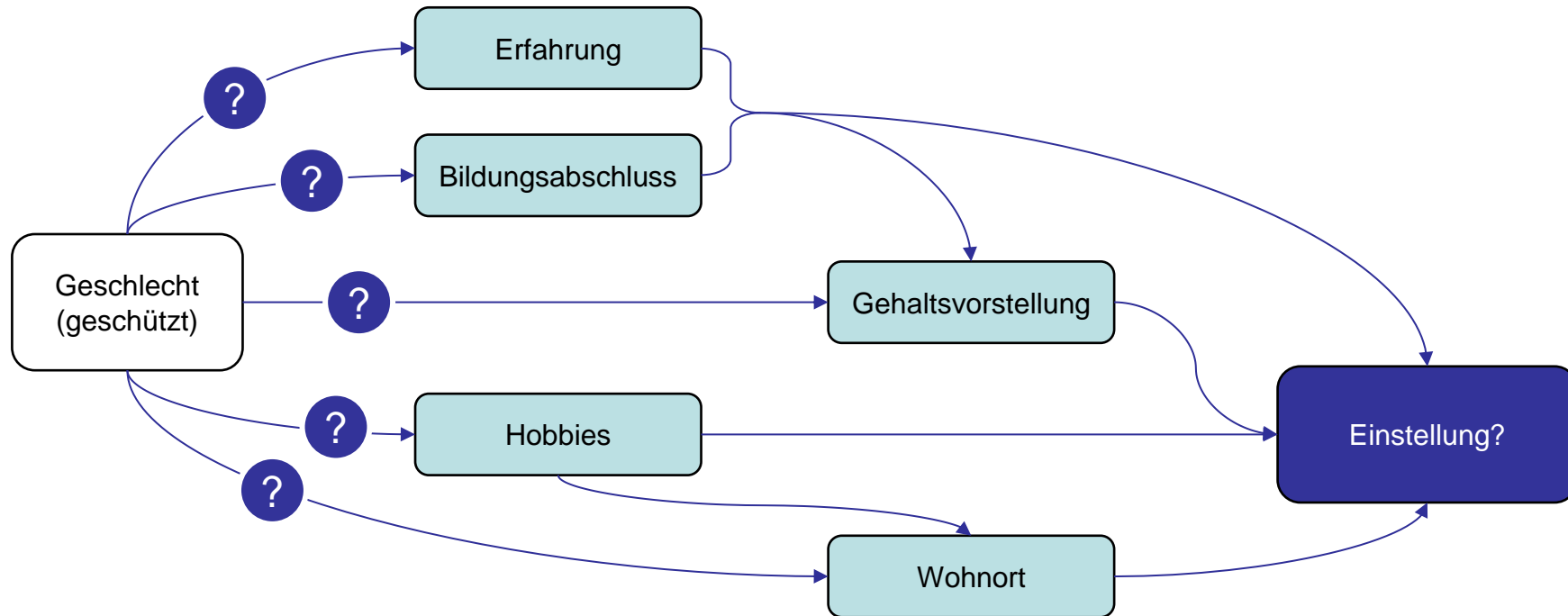
## Group Fairness

- ▶ Gleichbehandlung von Personengruppen
- ▶ Keine signifikanten Unterschiede nach Gruppen in der...
  - ▶ Leistungsfähigkeit des Modells
  - ▶ Verteilung der Ergebnisse



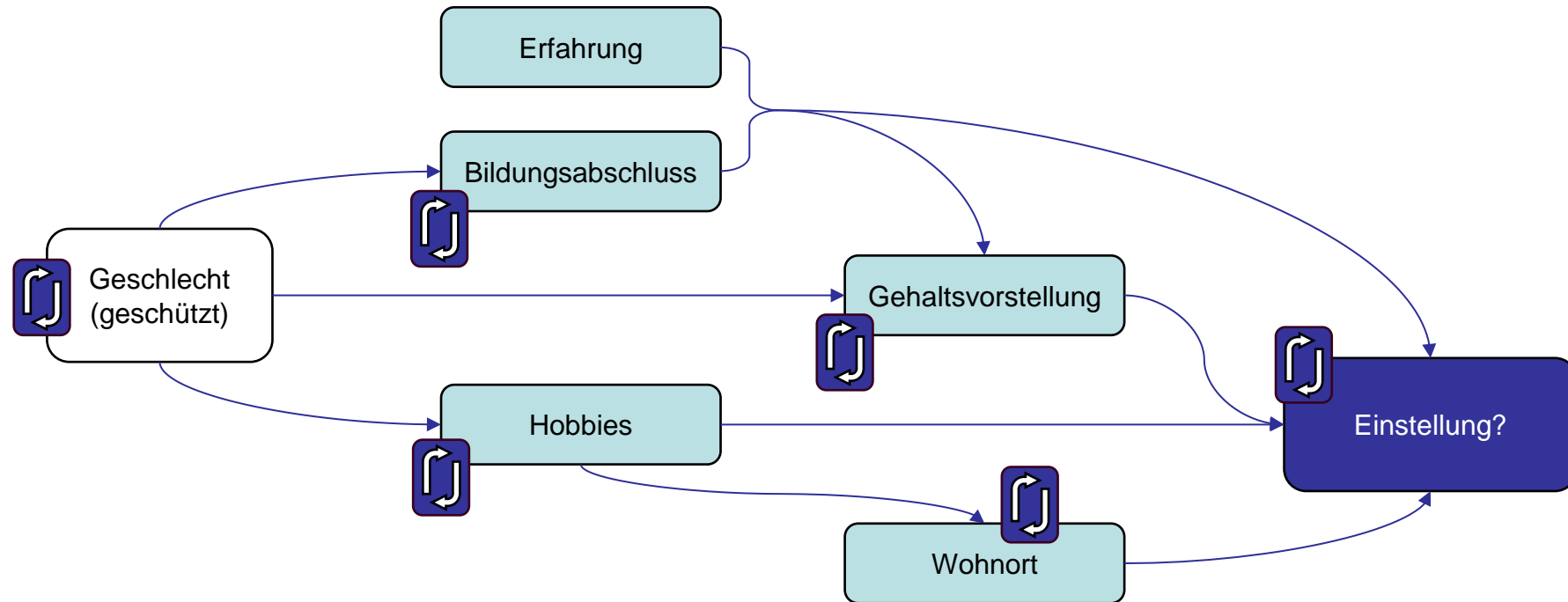
## Causal Fairness

- ▶ Analyse von kausalen Beziehungen zwischen Merkmalen und Ergebnissen
- ▶ Welchen Einfluss hat ein geschütztes Merkmal auf die Entscheidung?



## Counterfactual Fairness

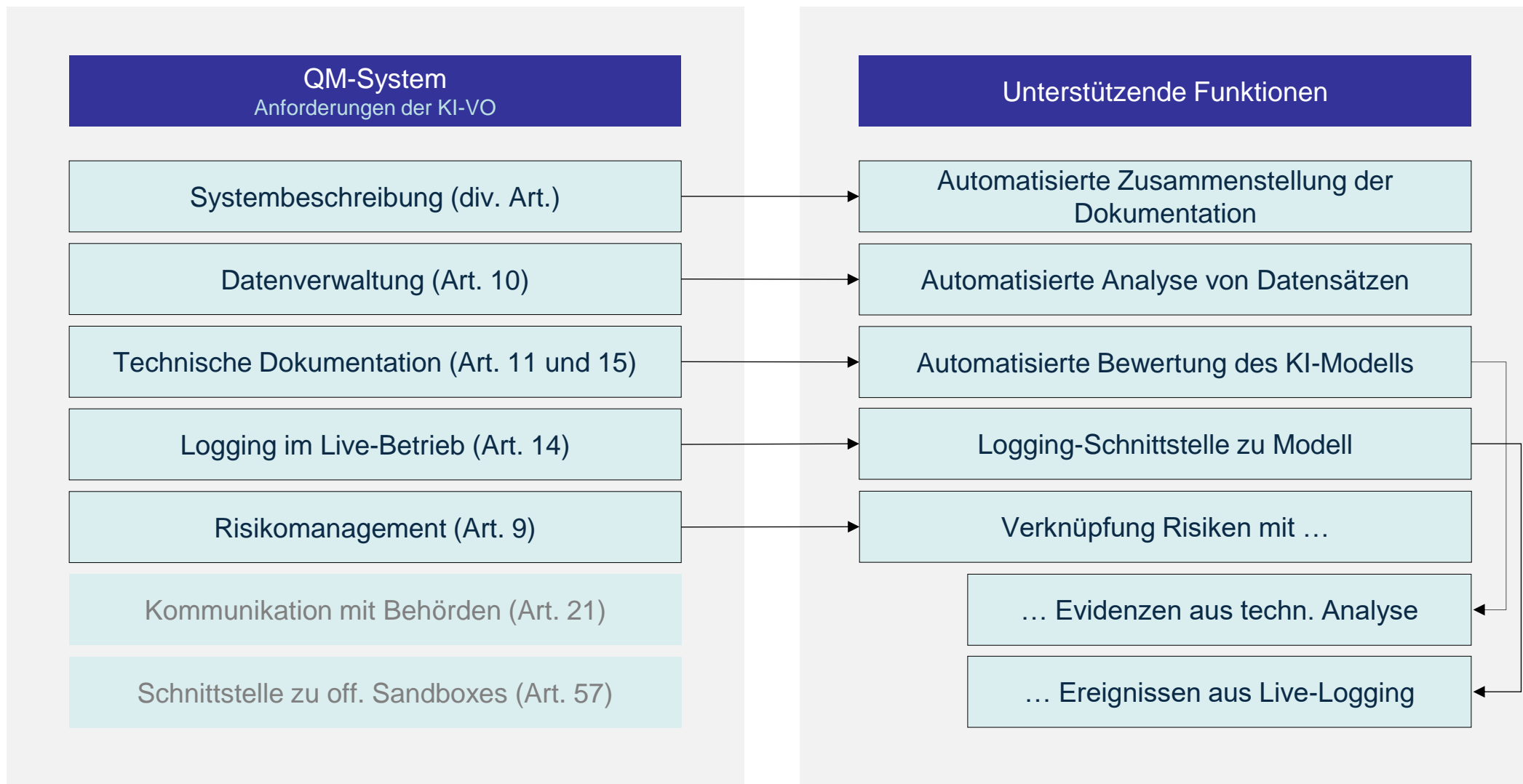
- ▶ Betrachtung kontrafaktischer Fragestellungen am kausalen Modell
- ▶ Untersuchung und Kompensation von Einflüssen geschützter Attribute





Art. 17 KI-VO sowie Anforderungen aus Kapitel III Abschnitt 2

# TECHNISCHE UNTERSTÜTZUNG FÜR QM-SYSTEME



Thetis Bewertung ^

Thetis Score ^

Daten zu Ihrer KI-Lösung ✓

Ergebnisse ○

Datenerfassung v

Konformitätsbewertung v

KI-Lösung schließen

[Startseite](#) > [KI-Lösung-Übersicht](#) > ... > [Ergebnisse](#)

## Ergebnisse

Wenn Sie neue [Daten zu Ihrer KI-Lösung](#) hochladen, kann der Thetis Score neu berechnet werden.

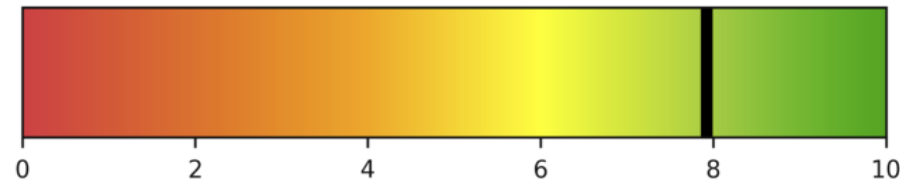
### Thetis Score

Datenqualität	Fairness	Qualität der Unsicherheit	Robustheit
<span style="color: red;">●</span> Schlecht	<span style="color: orange;">●</span> Mittel	<span style="color: green;">●</span> Gut	<span style="color: green;">●</span> Gut

Datenqualität Fairness Qualität der Unsicherheit Robustheit

### 2. Fairness-Auswertung

#### 2.1 Bewertung



Fairness-Auswertung Score (von 0 bis 10): 7.92

Fairness-Auswertung Bewertung: Mittel

#### 2.2 Ergebnisse und Empfehlungen

Der Klassifikator weist erhebliche Unterschiede bzgl. korrekter Prädiktionen für verschiedene Gruppen bestimmter sensibler Attribute auf, was auf signifikante Unterschiede in der Vorhersagegenauigkeit hindeutet.

Der Klassifikator zeigt ein ausgeglichenes Verhältnis bei der korrekten Identifizierung von Samples/Instanzen über verschiedene Gruppen hinweg für bestimmte sensitive Attribute, was auf faire Erkennungsraten zwischen verschiedenen Gruppen bestimmter sensibler



## Wissenswertes

Klicken Sie auf , um wissenswerte Informationen zum Element anzuzeigen.

Thetis Bewertung

Thetis Score

Daten zu Ihrer KI-Lösung

Ergebnisse

Datenerfassung

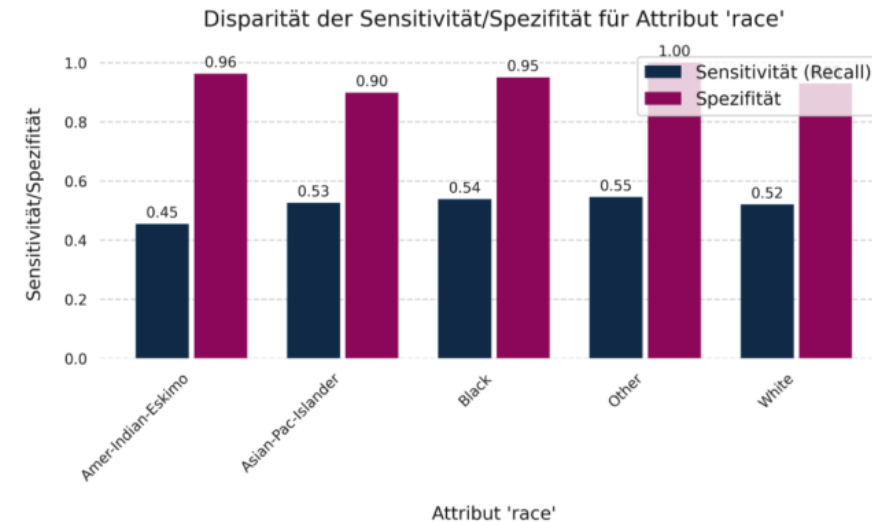
Konformitätsbewertung

KI-Lösung schließen

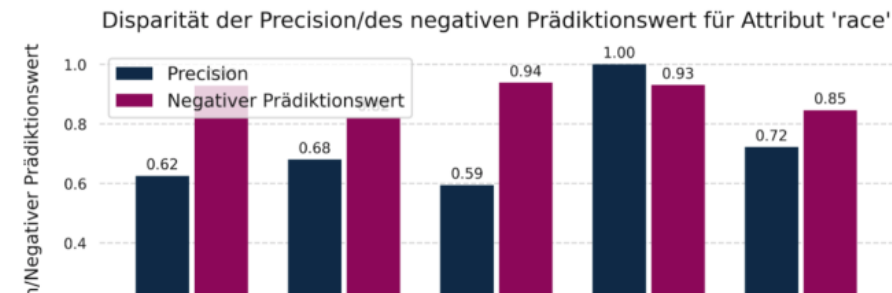
[Startseite](#) > [KI-Lösung-Übersicht](#) > ... > [Ergebnisse](#)

Attribute nicht Teil der eigentlichen Prädiktion sind und keinen Einfluss auf den Zielwert haben.

[Detaillierte Auswertungsergebnisse ausblenden](#)



**Abbildung A.4.1:** Disparität der Sensitivität (Recall) und Spezifität zwischen den Labels des sensiblen Attributs 'race'. Dieses Diagramm gibt Aufschluss darüber, inwiefern der Klassifikator in der Lage ist, Instanzen für verschiedene Labels korrekt zu identifizieren.



Wissenswertes



Helping you build ethical, fair and robust AI.

<https://thetis.de>

[thetis@efs-techhub.com](mailto:thetis@efs-techhub.com)

