



DIGITALE

GESELLSCHAFT

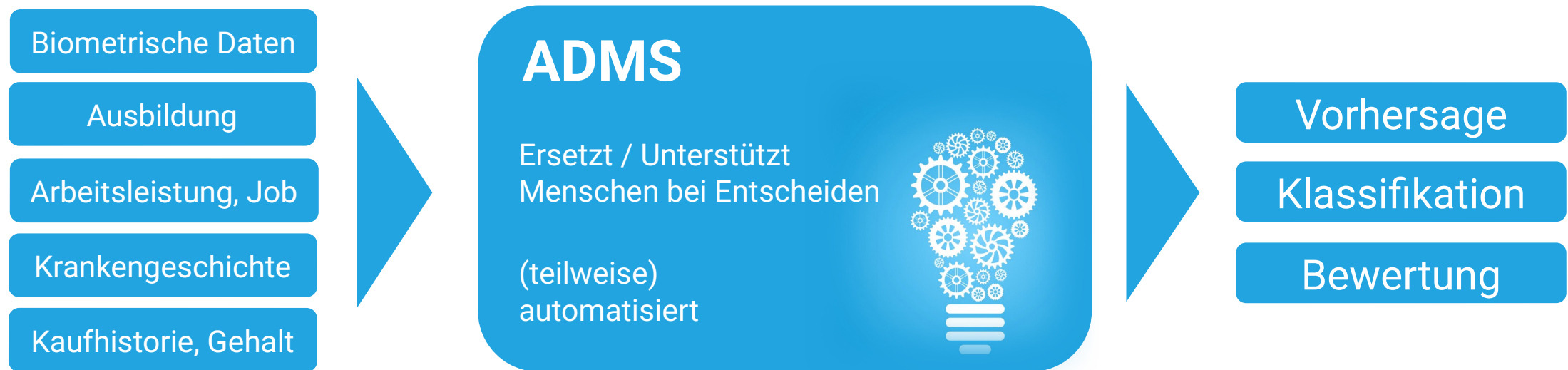
Regulierung künstlicher Intelligenz: Schweiz und International

Dr. Andreas Geppert | Dr. David Sommer

25. Februar 2023

Zürich - Volkshaus

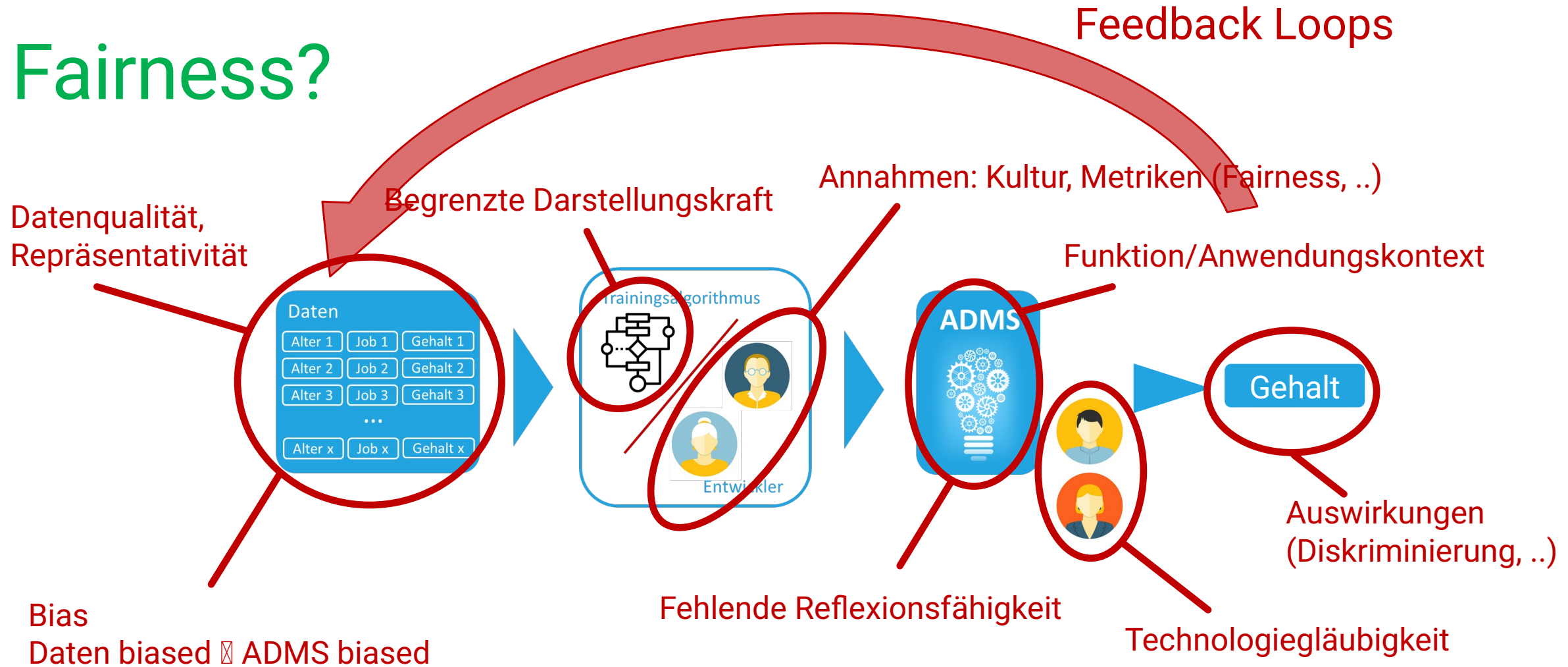
Automated Decision-Making Systems (ADMS)



ADM-Systeme: Entscheide zu sensiblen Lebensbereichen

Bsp.: Ausbildungsmöglichkeiten, Gesundheitsergebnisse, Arbeitsleistung, Mobilität, Interessen, Verhalten, persönliche Autonomie..

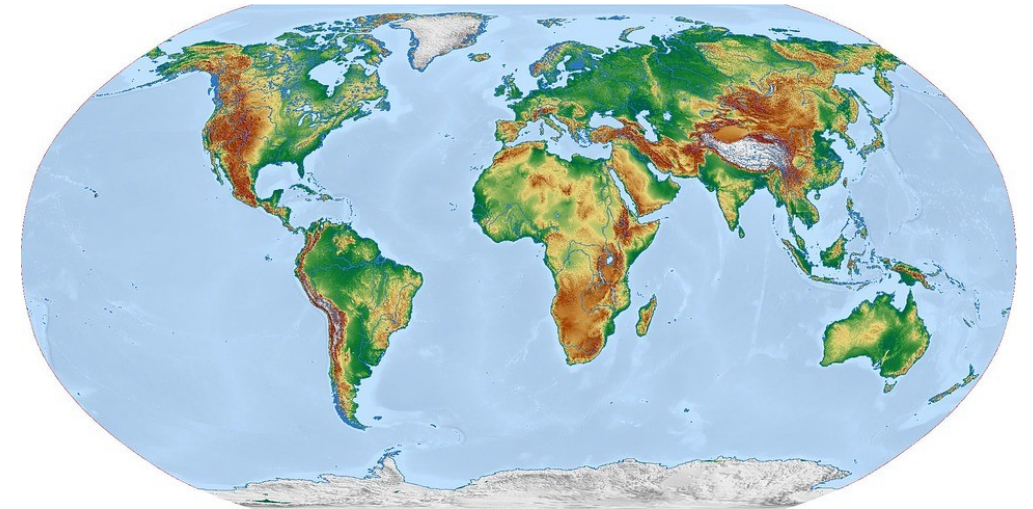
Fairness?



ADM-Systeme reproduzieren gesellschaftliche Normen

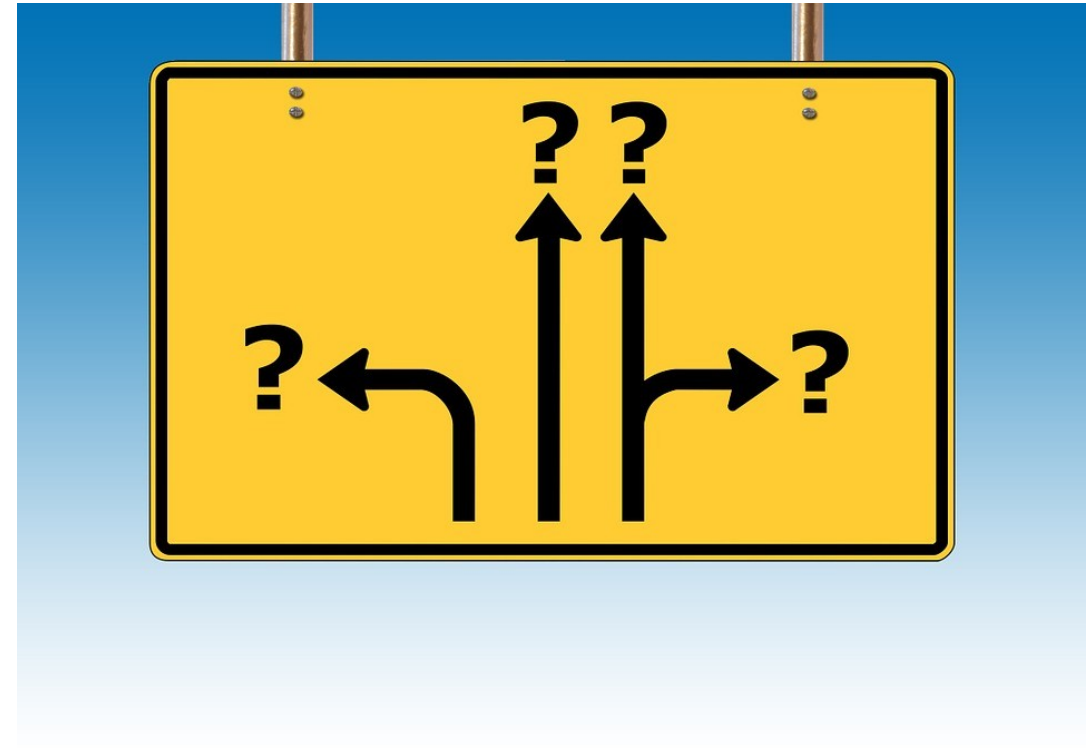
“Regulierungsvorschläge”

- AI Act
- Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (Council of Europe)
- Algorithmic Accountability Act of 2022 (USA)
- Ethical Norms for New Generation Artificial Intelligence (PRC)
- Principles for Responsible Algorithmic Systems (ACM)



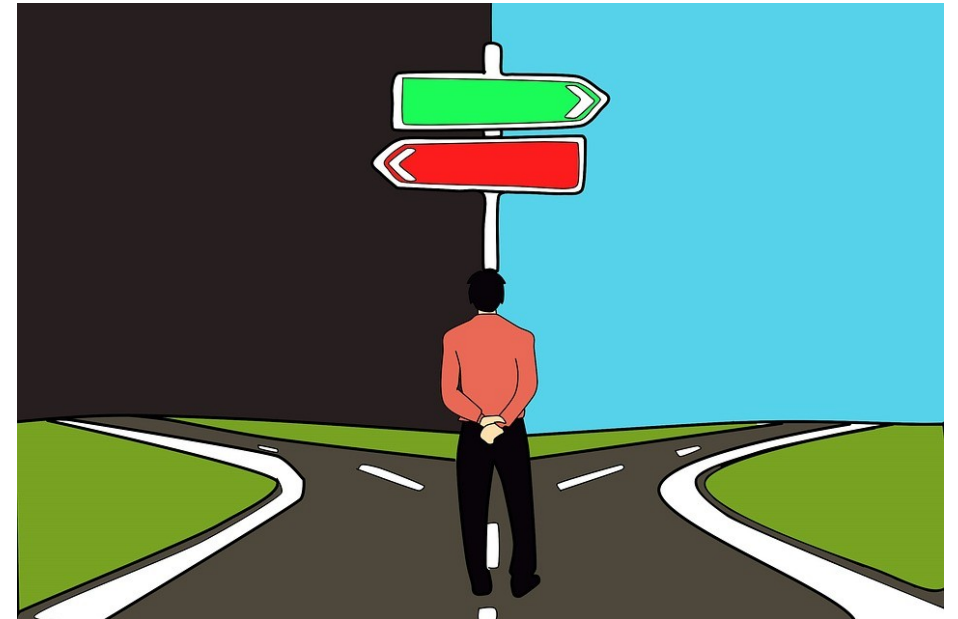
ADMS-Regulierung: Gemeinsamkeiten & Unterschiede

- Der Regulierungsbedarf wird generell anerkannt
- Scope: KI, Algorithmen, automatisierte Entscheidungen
- Automatisierte Entscheidungen als spezieller Problembereich
- Ansatz: risikobasiert, schadensbasiert
- Folgenabschätzung



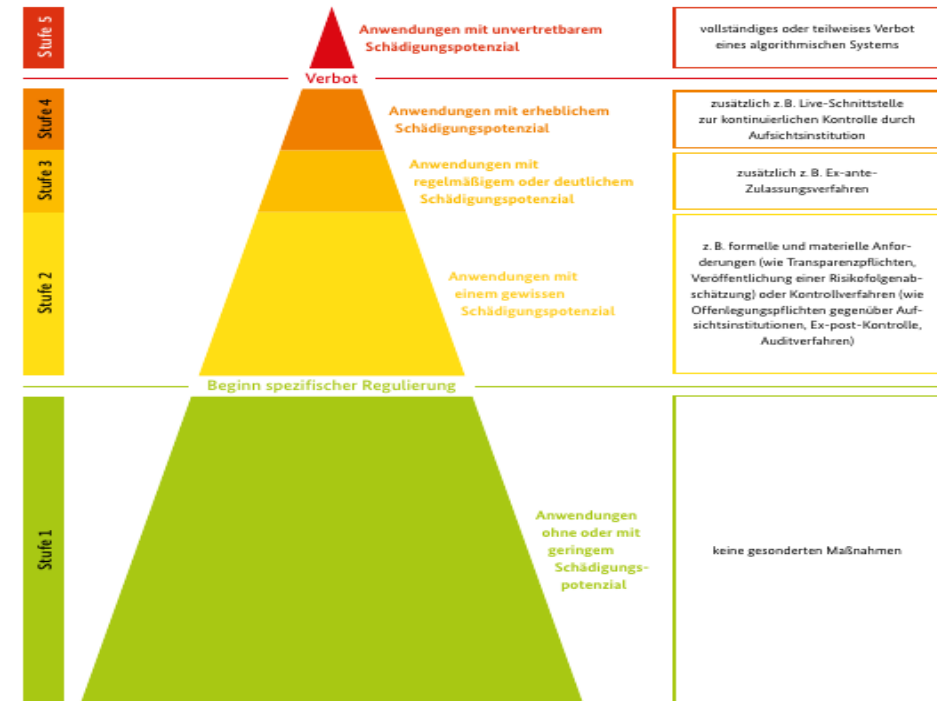
ADMS-Regulierung: Gemeinsamkeiten & Unterschiede (2)

- Gretchenfrage: Geltugsbereich
 - fallen staatliche Akteure (insb. Im Bereich staatliche Massenüberwachung) unter die Regulierung?
- Transparenz:
 - Verwendung von KI generell
 - Verwendete Algorithmen, Ansätze, Parameter, Koeffizienten, Trainingsdaten
- Menschliche Aufsicht
- Aufsichtsbehörde
- Systemverzeichnis



EU AI Act

- Gesetzesvorschlag der EU-Kommission
- Zu künstlicher Intelligenz allgemein
- Gilt für kommerzielle wie staatliche Akteure
- Risikobasierter Ansatz
 - Unakzeptables Risiko (verboten)
 - Hohes Risiko (Potentieller Schaden für Sicherheit, Gesundheit oder Grundrechte)
 - Niedriges oder kein Risiko



EU AI Act

- Risikomanagement-System
- Daten und Datenverwaltung
- Technische Dokumentation
- Protokollierung (Logging)
- Transparenz und Information der Nutzer
- Menschliche Aufsicht
- Genauigkeit, Robustheit, Sicherheit
- Konformitätsbewertung
- Registrierung in der EU-Datenbank

Algorithmic Accountability Act

- Gesetzesvorschlag im US-Repräsentantenhaus
- Kritische Entscheidungen über Konsument:innen (Bildung, Arbeit, Gesundheitswesen,...)
- Verlangt Folgenabschätzungen
- Federal Trade Commission als Aufsichtsbehörde
- Minimierung der negativen Auswirkungen
- Dokumentation der Konsument:innenrechte:
 - Kennzeichnung des ADMS
 - Opt-out-Möglichkeiten
 - Widerspruchs- und Korrekturmöglichkeiten
- Öffentlich zugängliches Verzeichnis der Zusammenfassungen

Ethical Norms for New Generation AI

- Publiziert vom National New Generation Artificial Intelligence Governance Specialist Committee
- Zweck: ethische Richtlinien für juristische Personen und andere Institutionen im Bereich KI
- 6 grundlegende ethische Normen / 18 Richtlinien:
 - Förderung des menschlichen Wohlergehens
 - Förderung von Fairness und Gerechtigkeit
 - Schutz der Privatsphäre und der Sicherheit
 - Sicherung der Kontrolle und des Vertrauens
 - Stärkung der Verantwortlichkeit
 - Verbesserung der Kulturethik

Stmnt on Principles f. Responsible Algorithmic Systems

- Publiziert vom US/European Technology Policy Committee der ACM
- 9 Prinzipien
- Risiken verhältnismässig zum adressierten Problem
- Schadensminimierung (Folgenabschätzung)
- Anforderungen an Entwickler:innen und Entwicklung
- Transparenz
 - Wie wurden Datensets, Variablen und Modelle ausgewählt
 - Angabe des Konfidenzlevels
- Unabhängige Verifizierung und Validierung bei lebenswichtigen Entscheidungen
- Widerspruchsmöglichkeiten und Auditierbarkeit

Unsere Forderungen

- Warum wollen wir regulieren?
 - **Vertrauen** und **Rechtsicherheit**
- Was wollen wir regulieren?
 - Was ist das Problem?
 - **Definition** von ADMS?
- Wie können wir wissen, wo und wie ADMS im Einsatz ist?
 - **Transparenz**
 - Welche Art und auf welcher Ebene?
- Wie können wir “**Ungerechtigkeit**” vermeiden oder ausgleichen?
 - Ordentliche ML-Entwicklungsprozesse
 - Risikoklassen
 - **Verantwortlichkeit und Haftung garantieren, Verbote und Moratorien verhängen**
- ...gleichzeitig Innovationen fördern und bürokratischen Aufwand begrenzen



Vorschlag der
Digitalen
Gesellschaft

Ein ADM-Gesetz

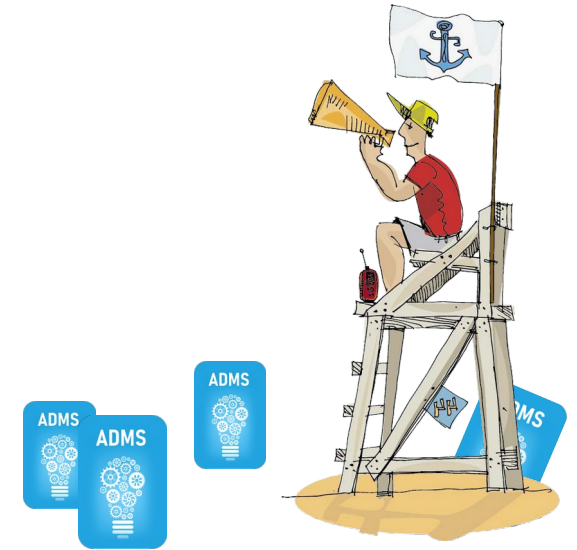
- Die **Digitale Gesellschaft** fordert
- **Transparenz** beim Einsatz der Systeme
 - Für Staat und Privatwirtschaft
- **Erklärbarkeit**
 - Einzelne Systeme sowie Verkettung
 - Ziel: Ausreichend für grobe, externe Risikoeinschätzung
- **Ansatz**: Mischform zwischen Risiko- und Schadensbasiert
 - Risiko: präventive Massnahmen nach Risiko (negligible, high, forbidden)
 - Schaden: Vergeltung retrospektiv nach Schadensfall



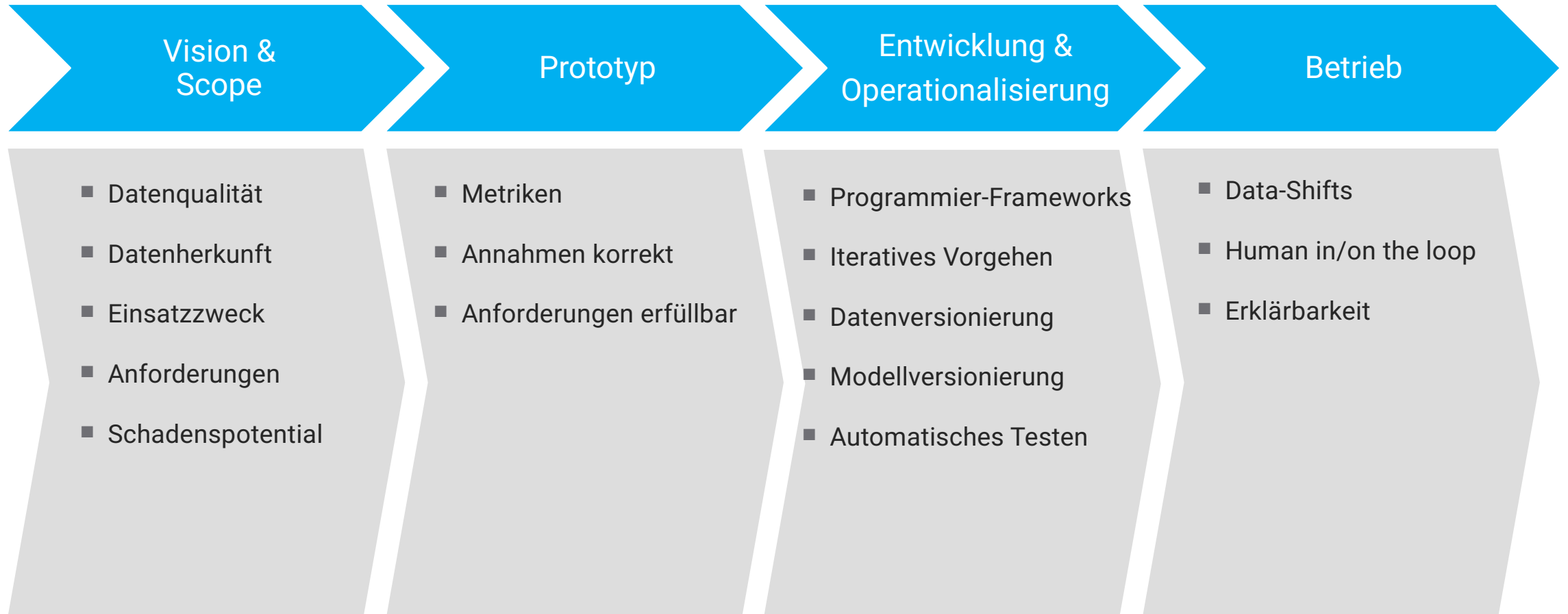
Details, Begründung und
Einschätzungshilfen:
Positionspapier

Unterstützung und Sanktionen

- Die ADM-Aufsicht
 - Staatliches Kompetenzzentrum
 - Für Unternehmen, Behörden und die Öffentlichkeit
 - Kontrolliert Gesetzeskonformität
 - Erstellt Checklisten und Good-Practices-Anleitungen
 - Orchestriert Langzeitanalysen
- **Verbandsklagerecht**
 - Verbände klagen ohne persönliche Betroffenheit
- **Beweislastumkehr**
- Umsatzabhängige **Verwaltungssanktionen**



ML-Entwicklungsprozesse



Ausblick

- Aktivitäten auf diversen politischen Ebenen
 - kantonal, national, supranational
- Tendenz zu einem risikobasierten Ansatz
 - EU ex ante, Schweiz eher ex post
- Tendenz zu Zertifizierung
 - Nutzen fraglich
- ADM Aufsicht
 - Ressourcen und Kompetenzen nötig
 - wirksame Massnahmen, nicht nur Compliance



Bild-Quelle: Bildsuche Powerpoint, freie Lizenz

➡ Positionspapier der Digitalen Gesellschaft
www.digitale-gesellschaft.ch