

KI und Ethik

Die Ethischen Leitlinien für Künstliche Intelligenz der EU Kommission und ähnliche Initiativen

Christof Höfner
Senior Legal Counsel
Zurich Corporate Center



KI und Ethik

- Die AI HLEG Ethic Guidelines für vertrauenswürdige künstliche Intelligenz
- Die OECD Prinzipien zu Künstlicher Intelligenz
- Das «KI Qualitätssiegel» (KI Bundesverband e.V.)
- Telia «Guiding Principles on Trusted AI Ethics»
- Das deutsche "KI-Observatorium,, des BMAS

Hochrangige EU Expertengruppe für künstliche Intelligenz “AI HLEG”

- 2018 ernannte die EU Kommission 52 Experten zur «High-Level Expert Group on AI»
- **Ziel der KI Expertengruppe :**
 - Förderung und Einführung einer Europäischen KI Strategie
 - Erarbeitung von Vorschlägen für zukunftsorientiertes und ethisches policy development, **rechtliche und gesellschaftliche AI Themen**, sozial-wirtschaftliche Herausforderungen
 - Die AI HLEG dient als Lenkungsausschuss für die Arbeit der Europ. KI Allianz
- **Die “KI Ethikleitlinien”:**
 - Dezember 2018: die AI HLEG präsentierte den ersten Entwurf der «AI Ethics Guidelines»
 - Die Leitlinien waren bis 18. Januar 2019 zur Konsultation offen und wurden am 9. April veröffentlicht

<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>
<https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

EU High-Level Expert Group on Artificial Intelligence

Kapitel I: Basis vertrauenswürdiger KI

Basis: die Grundrechte

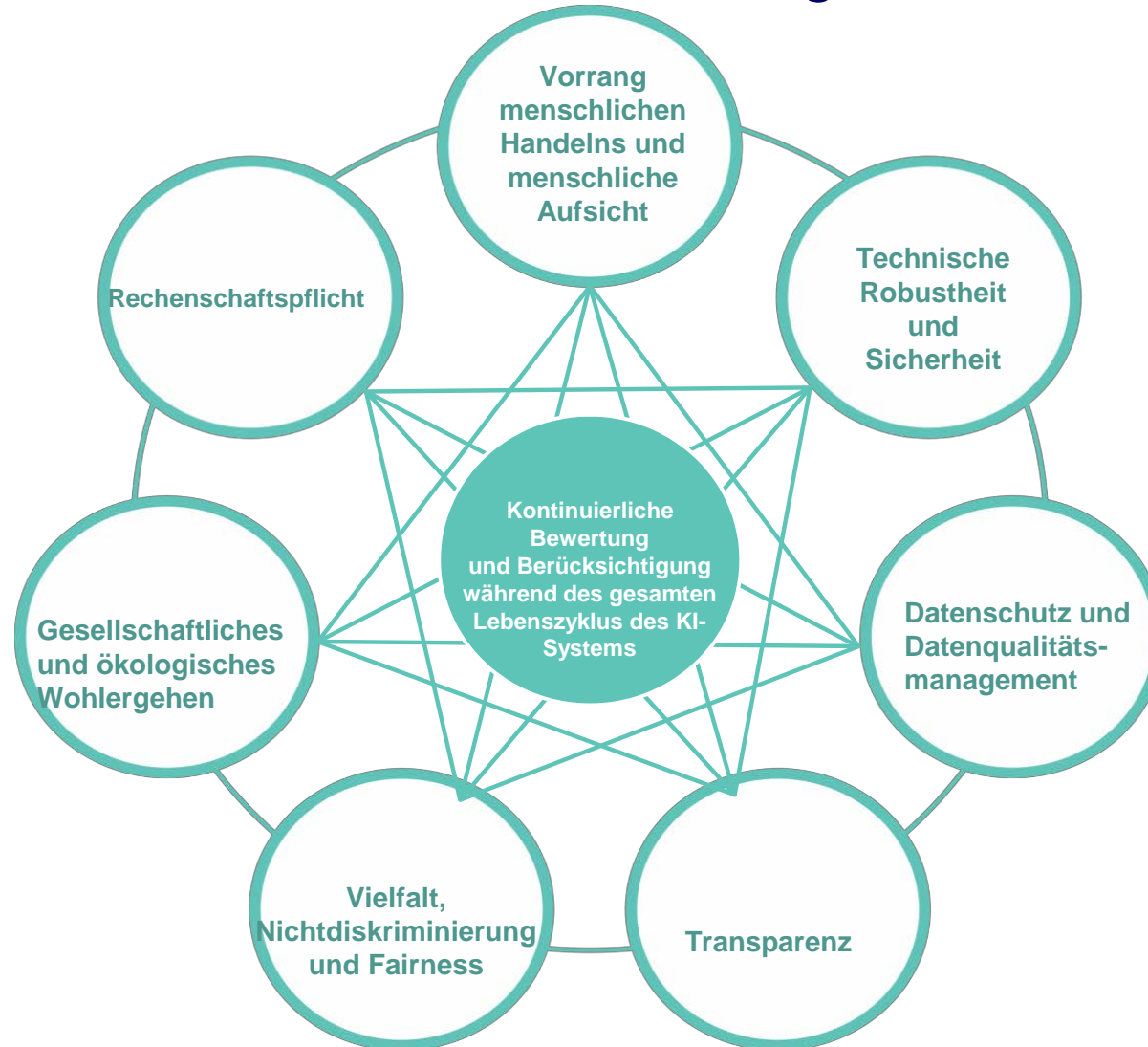
- Achtung der Menschenwürde
- Freiheit des Einzelnen
- Achtung vor Demokratie, Gerechtigkeit und Rechtsstaatlichkeit
- Gleichheit, Nichtdiskriminierung und Solidarität
- Bürgerrechte

Ethische Grundsätze:

- (i) Achtung menschlicher Autonomie
- (ii) Schadensverhütung
- (iii) Fairness
- (iv) Erklärbarkeit

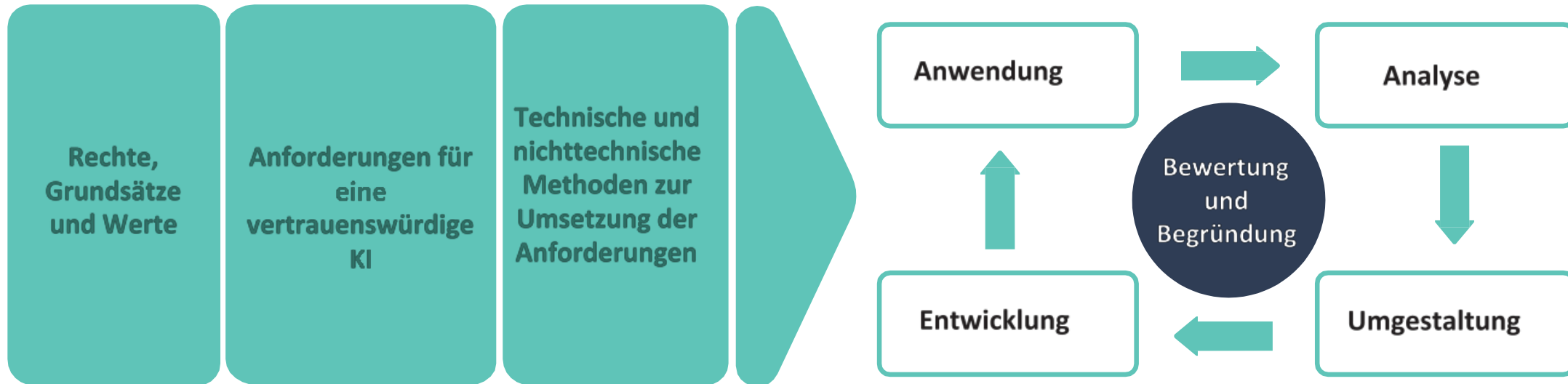
EU High-Level Expert Group on Artificial Intelligence "AI HLEG"

Kapitel II (1) Fundamente einer vertrauenswürdigen KI:



EU High-Level Expert Group on Artificial Intelligence “AI HLEG”

Kapitel II (2): Verwirklichung einer vertrauenswürdigen KI



EU High-Level Expert Group on Artificial Intelligence

Kapitel III (1): Bewertungsliste für die praktische Anwendung vertrauenswürdiger KI (Pilot)

1) Vorrang menschlichen Handelns und menschlicher Aufsicht

- Grundrechte
- Vorrang menschlichen Handelns
- Menschliche Aufsicht

2) Technische Robustheit und Sicherheit

- Schutz gegen Angriffe und Sicherheit
- Rückfallplan und allgemeine Sicherheit
- Präzision
- Zuverlässigkeit und Wiederholbarkeit

EU High-Level Expert Group on Artificial Intelligence

Kapitel III (2): Bewertungsliste für die praktische Anwendung vertrauenswürdiger KI (Pilot)

3) Schutz der Privatsphäre und Datenqualitätsmanagement

- Achtung der Privatsphäre und Gewährleistung des Datenschutzes
- Qualität und Integrität der Daten
- Datenzugriff

4) Transparenz

- Rückverfolgbarkeit
- Erklärbarkeit
- Kommunikation

EU High-Level Expert Group on Artificial Intelligence

Kapitel III (3): Bewertungsliste für die praktische Anwendung vertrauenswürdiger KI (Pilot)

5) Vielfalt, Nichtdiskriminierung und Fairness

- Vermeidung unfairer Verzerrungen
- Barrierefreiheit und universeller Entwurf
- Beteiligung von Interessengruppen

6) Gesellschaftliches und ökologisches Wohlergehen

- Nachhaltige und umweltfreundliche KI
- Soziale Auswirkungen
- Gesellschaft und Demokratie

EU High-Level Expert Group on Artificial Intelligence

Kapitel III (4): Bewertungsliste für die praktische Anwendung vertrauenswürdiger KI (Pilot)

7) Rechenschaftspflicht

- Nachprüfbarkeit
- Minimierung und Meldung negativer Auswirkungen
- Dokumentation von Kompromissen
- Rechtsschutz

Die HLEG KI bat alle Beteiligten die Bewertungsliste zu testen.

Wir fordern alle Beteiligten auf, diese Bewertungsliste in der Praxis zu erproben und uns eine Rückmeldung über deren Umsetzbarkeit, Vollständigkeit und Relevanz für die spezifische KI-Anwendung oder den Anwendungsbereich sowie über Überschneidungen oder Entsprechungen mit bestehenden Konformitäts- oder Bewertungsverfahren zu geben ..

The OECD Principles on Artificial Intelligence

“Eine Empfehlung”

Die OECD-Grundsätze für künstliche Intelligenz wurden am 22. Mai 2019 von den Mitgliedstaaten verabschiedet.

- 1) KI soll den Menschen und dem Planeten zugute kommen, indem sie integratives Wachstum, nachhaltige Entwicklung und Wohlstand fördert.
- 2) KI-Systeme sollen so konzipiert sein, dass sie die Rechtsstaatlichkeit, die Menschenrechte, die demokratischen Werte und die Vielfalt respektieren, und angemessene Garantien enthalten - zum Beispiel, die es ermöglichen, dass Menschen ggf. eingreifen -, um eine gerechte und gerechte Gesellschaft zu gewährleisten.
- 3) Es soll Transparenz und verantwortungsvolle Offenlegung in Bezug auf KI-Systeme geben, damit die Menschen die KI-basierten Ergebnisse verstehen und in Frage stellen können.
- 4) KI-Systeme müssen während ihres gesamten Lebenszyklus robust, sicher und zuverlässig funktionieren, und potenzielle Risiken sollten kontinuierlich bewertet und verwaltet werden.
- 5) Organisationen und Einzelpersonen, die KI-Systeme entwickeln, einsetzen oder betreiben, sollten für ihr ordnungsgemäßes Funktionieren im Einklang mit diesen Grundsätzen verantwortlich gemacht werden.

Andere Ethik Guidelines für künstliche Intelligenz (3)

KI Bundesverband e.V. «AI-Qualitätssiegel»

- KI Bundesverband e.V.
160+ innovative Unternehmen, SMEs, start-ups und Experten
<https://ki-verband.de/>
- AI-Qualitätssiegel-Kriterien (als Selbstverpflichtung zu bestätigen):
 1. Ethik
 2. Unparteilichkeit
 3. Transparenz
 4. Sicherheit und Datenschutz

© KI Bundesverband e.V.



Andere Ethik Guidelines für künstliche Intelligenz (4)

Telia «Guiding Principles on Trusted AI Ethics»



© Telia Company AB

<https://www.teliacompany.com/globalassets/telia-company/documents/about-telia-company/public-policy/2018/guiding-principles-on-trusted-ai-ethics.pdf>

TÜV für künstliche Intelligenz kommt

SZ 12.11.2019

- Noch in diesem Jahr soll nach SZ-Infos ein "deutsches KI-Observatorium" die Arbeit aufnehmen. Eine Art [TÜV](#) für künstliche Intelligenz in Unternehmen. Anfang 2020 ist die offizielle Eröffnung durch Bundesarbeitsminister Heil geplant.
- Das KI-Observatorium soll als Einheit im Bundesministerium für [Arbeit und Soziales](#) (BMAS) angesiedelt werden. Erste Beschäftigte werden bereits eingestellt.
- Mittelfristig ist geplant, ein eigenes Bundesinstitut für KI mit deutlich mehr Personal einzurichten. Es soll Chancen und Risiken der künstlichen Intelligenz bewerten und politisch steuern helfen.
- https://www.sueddeutsche.de/wirtschaft/ki-observatorium-tuev-arbeitsministerium-1.4676937?sc_src=email_1131102&sc_lid=111041537&sc_uid=Ws5RRvr7C7&sc_lid=33064&utm_medium=email&utm_source=emarsys&utm_content=www.sueddeutsche.de%2Fwirtschaft%2Fki-observatorium-tuev-arbeitsministerium-1.4676937&utm_campaign=Espresso+am+Morgen+12.11.2019

“Our Machines should be nothing more than tools for extending the power of human beings who use them.”

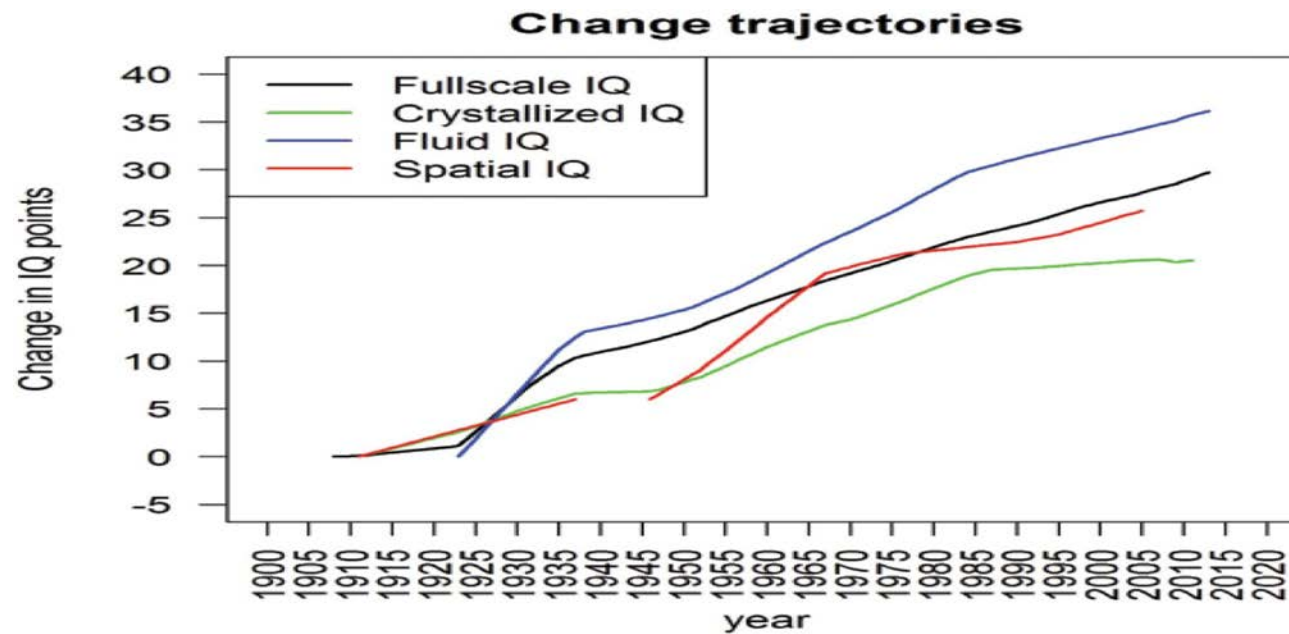
Thomas J. Watson, Jr. - IBM President and CEO 1956 - 1971

Christof Höfner
Senior Legal Counsel
Zurich Corporate Center
+41 767485869

Natürliche Intelligenz

"Flynn Effect"

Der Flynn-Effekt beschreibt den erheblichen und lang anhaltenden Anstieg der Intelligenztestergebnisse, die im Laufe des 20. Jahrhunderts in vielen Teilen der Welt gemessen wurden.



Natürliche Intelligenz

Der negative Flynn Effect

In jüngster Zeit scheint es Anzeichen dafür zu geben, dass sich der Flynn-Effekt seit Ende 1990 umgekehrt hat. Es wurden einige Tests durchgeführt, die zeigten in wie vielen Ländern 1) das umgekehrte Phänomen inzwischen aufgetreten ist. 2)

Es scheint eine gewisse Korrelation zwischen Intelligenz und Schulsystemen, Lernmethoden, Lesen von Büchern, Einsatz von IT und Phantasie zu bestehen.

Professor James Flynn sagte in einem Interview:

"Stellen Sie sicher, dass Ihre Kinder eine Leidenschaft für Ideen haben.

Der genetische Unterschied in der Gehirnphysiologie könnte bis 25 Prozent zu IQ-Unterschieden beitragen. Gene sind wichtig, aber das Haben und der Glaube an die eigenen Ideen ist ebenso wichtig.... " 3)

- 1) Thomas Teasdale: Denmark, Germany, Austria, Switzerland, France, UK
- 2) Edward Dutton, Dimitri van der Linden, Richard Lynn (Science Direct 2016)
- 3) <http://www.zeit.de/zeit-wissen/2008/02/Flynn-Interview/komplettansicht>