

# VERTAUENSANKER UND MITBESTIMMUNG

WAS AUFSICHTSRÄTE UND ARBEITSDIREKTOREN TUN  
KÖNNEN UM KI AM MENSCHEN AUSZURICHTEN

DÜSSELDORF 7. SEPTEMBER 2022

LOTHAR SCHRÖDER



# VERTRAUEN IST DER MÖRTEL DER WIRTSCHAFT UND GESELLSCHAFT ZUSAMMENHÄLT, DIESER MÖRTEL BRÖCKELT

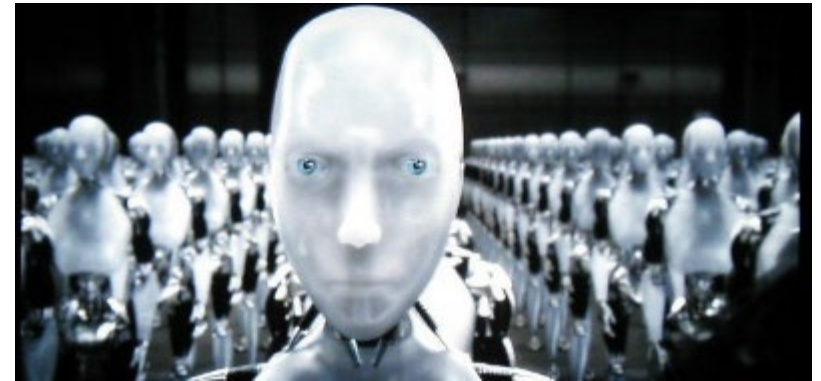
Vertrauensseligkeit war gestern:

- Fake news
- Abgasskandale
- Missbrauchsfälle in der Katholischen Kirche
- Krieg und Desinformation
- Halbwertzeit von Coronastrategien

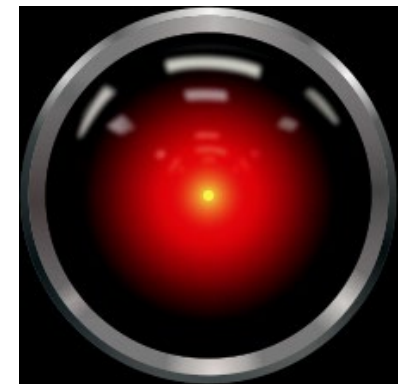
„Vertrauen ist der Wille, sich verletzlich zu zeigen.“

M. Osterloh, A. Weibel: *Investition Vertrauen. Prozesse der Vertrauensentwicklung in Organisationen.* Gabler, Wiesbaden 2006.

# VERTRAUEN IN MASCHINELLE ENTSCHEIDUNGEN? UNSERE BILDER VON KI



„Betriebliche Mitbestimmung und eine frühzeitige Einbindung der Betriebsräte stärken das Vertrauen und die Akzeptanz der Beschäftigten bei der Einführung und der Anwendung von KI „

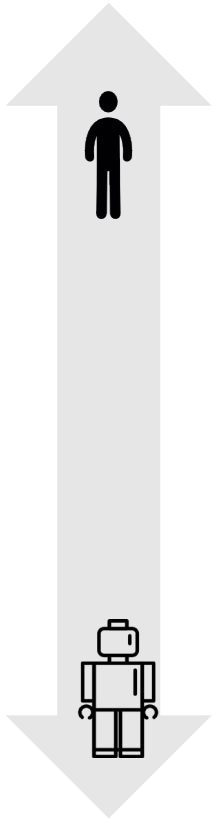


KI Strategie Bundesregierung Seite 28

# SCHAFFT BISHERIGE DIGITALISIERUNGSERFAHRUNG VERTRAUEN ? SORGEN UND ÄNGSTE VON BESCHÄFTIGTEN

- Kontrollverlust zu erleiden und Übersicht zu verlieren
- mit steigender Veränderungsgeschwindigkeit nicht mithalten zu können
- den Job / die Karrierechancen zu verlieren
- Informationelle Selbstbestimmung abzugeben
- Qualifikationen entwertet zu bekommen
- Diskriminierung zu erfahren
- steigendem Arbeitsdruck ausgesetzt zu werden
- einer Entmenschlichung ausgeliefert zu sein

# UNSERE VERTRAUENKONZEPTE BRINGEN ABSTUFUNGEN HERVOR - WIR ÜBERLASSEN SO MANCHES SCHON DEN MASCHINEN



1. **Der Mensch macht alles autonom**
2. **Der Computer bietet Alternativen an**
3. **Der Computer bietet eine Auswahl an Alternativen an**
4. **Der Computer schlägt eine Alternative vor**
5. **Der Computer führt eine Alternative nach Freigabe aus**
6. **Der Computer führt eine Alternative aus, der Mensch hat ein Veto**
7. **Der Computer führt eine Alternative aus und informiert den Menschen**
8. **Der Computer führt eine ausgewählte Alternative aus und informiert den Menschen nach Anforderung**
9. **Der Computer führt eine ausgewählte Alternative aus und informiert den Menschen nach maschineller Entscheidung**
10. **Der Computer handelt komplett autonom**

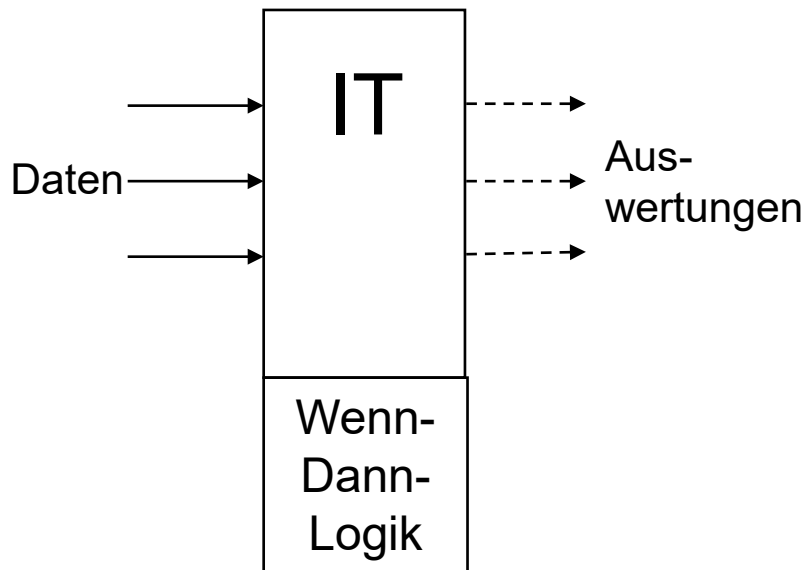
Quelle Miller & Parasuraman 2007

# DIE EIGENARTEN VON KI-SYSTEMEN VERLANGEN NACH EINER NEUEINSTELLUNG UNSERER VERTRAUENSKULTUR

- Erweiterte Optionen in die **Autonomie** des Menschen einzugreifen
  - **Sensorische Fähigkeiten** übersteigen sinnliche Wahrnehmung bei weitem
  - Lernfähigkeit erzeugt sich **verändernde Werkzeuge**
  - Bedeutungszuwachs für rationale Schlussfolgerungen
  - Auswirkung auf eine **Vielzahl von Persönlichkeitsrechten**
- **Vielfalt qualitätsrelevanter Normen, Vorgaben und Indikatoren**
  - Beschleunigungswirkung für betriebliche Prozesse
  - Fähigkeit zur Analyse und Verknüpfung auch **willentlich nicht erzeugter Daten**
  - Unterschiedliche **Kritikalitätsstufen**
  - Ordnungsrahmen muss auf partiell unvollständiger Information aufsetzen

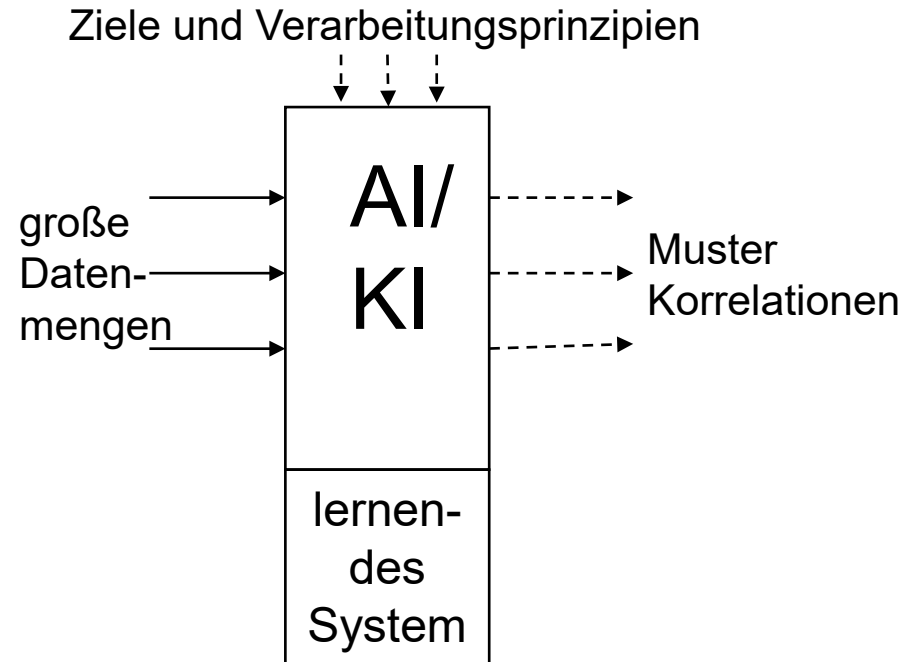
# Herkömmliche Gestaltungsmuster verlieren zum Teil an Wirksamkeit

## Herkömmliche Algorithmen



Regime des „Vorhersehbaren“

## Lernende Maschinen



Regime des „nicht Vorhersehbaren“

# PERSÖNLICHKEITSRECHTE BRAUCHEN MEHR ALS DIE REGLEMENTIERUNG DER LEISTUNGS- ODER VERHALTENSKONTROLLE

Merkmale und Anwendungen der KI	Persönlichkeitsrechte ( GrCH )	Mögliche Reglements
<ul style="list-style-type: none"><li>• Maschinelle Arbeitssteuerung</li><li>• Personalentscheidungen</li><li>• Sensorische Sprach – und Mimikauswertungen</li><li>• Lernen zur Optimierung</li><li>• Kommunikationsfähigkeit</li></ul>	<ul style="list-style-type: none"><li>• <b>Würdige und gesunde Arbeitsbedingungen</b></li><li>• <b>Chancengleichheit, Nichtdiskriminierung</b></li><li>• <b>Gedankenfreiheit</b></li><li>• <b>Demokratische Kontrolle und Soziale Verantwortung, Vielfalt</b></li><li>• <b>Freie Meinungsbildung</b></li></ul>	<ul style="list-style-type: none"><li>• Grenzen maschineller Autonomie</li><li>• Training ,Fairness und Genauigkeit, Widerspruchsrechte</li><li>• Reglementierung Einsatzfelder und Profilbildung</li><li>• Ständige Evaluation</li><li>• Transparenz, Alternativen</li></ul>



# KI UND TELEKOM – EINIGE SCHLAGLICHTER

- Beethovens vollendete 10 te Sinfonie
- KI- gestützte Netzausbauplanung FTTH
- KI-Ethik Leitlinien
- T-Labs : Mit KI gegen Hassbotschaften
- Frag Magenta
- T-Systems: KI Entwicklerplattform in der Cloud
- Deutsche Funkturm: Roboter schreiben Mietverträge
- DT Service: KI- Assistenz
- KI-im Kundenkontaktmanagement
- KI-macht in Entertain Filmempfehlungen
- „Hallo magenta“ lernt Sprache



Videsequenzen psychologisch ausgewertet:  
Offenheit:

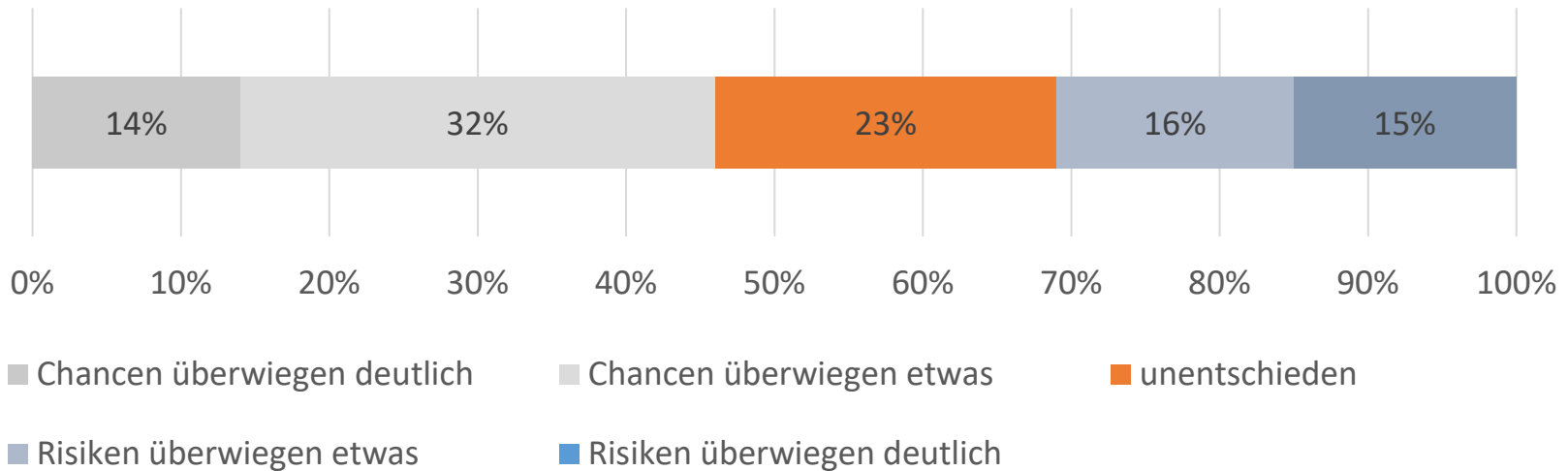
Sunday Pichai  
Google 78,7%  
Timotheus Höttges  
Telekom 74,9

Welt am Sonntag 16.09.20



# KI: CHANGE ODER RISIKO

Frage: Sehen Sie in Künstlicher Intelligenz eine Chance oder ein Risiko?

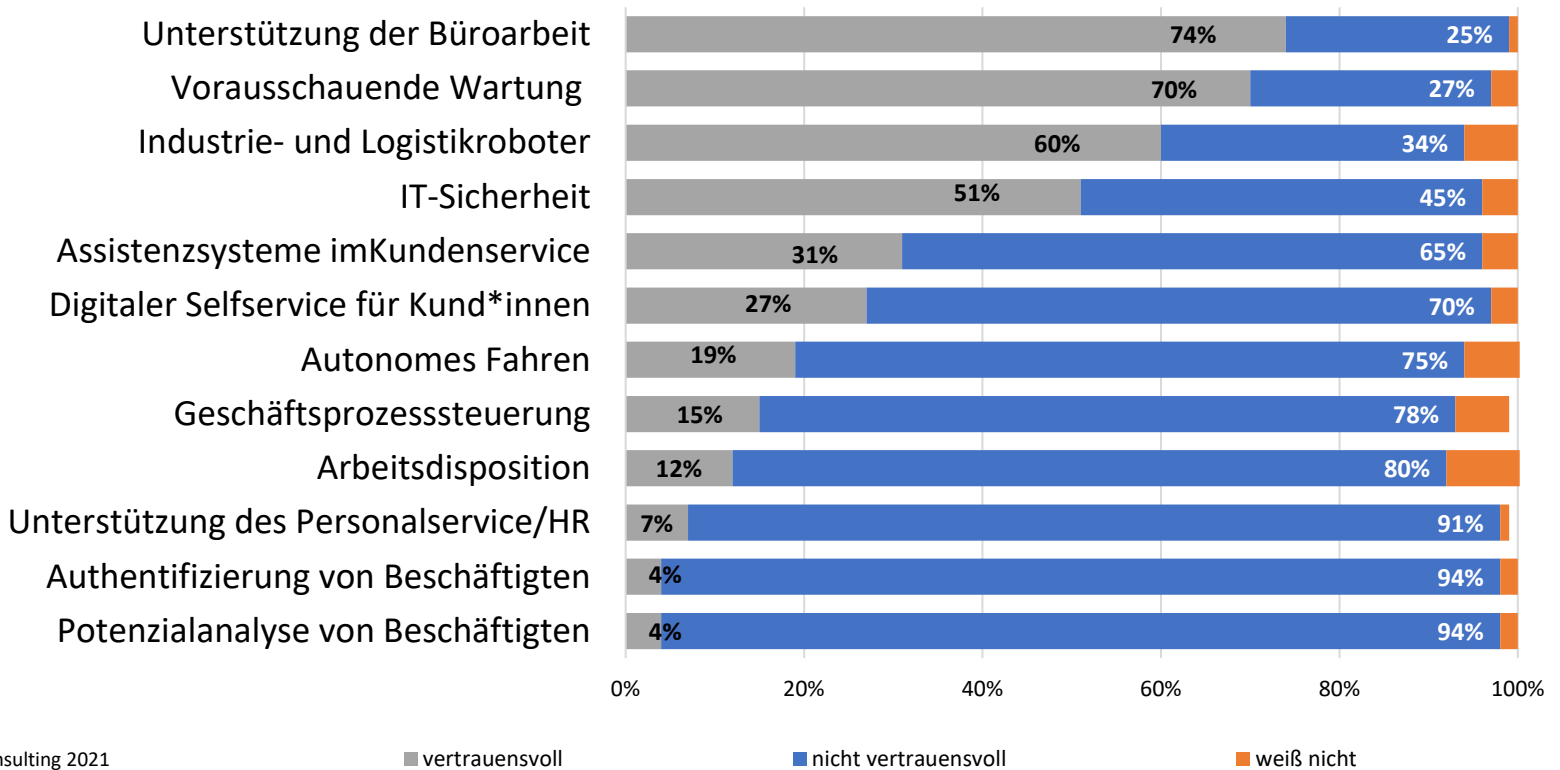


Quelle : Input- Consulting November 2021

# VERTRAUEN WOFÜR ?

## ANWENDUNGSSPEZIFISCHES VERTRAUEN VON BETRIEBS- UND PERSONALRÄTEN

Frage: Für wie vertrauensvoll halten Sie den Einsatz von KI-Systemen in unterschiedlichen beruflichen Anwendungsgebieten?  
(Anteil Antworten „vertrauensvoll“)

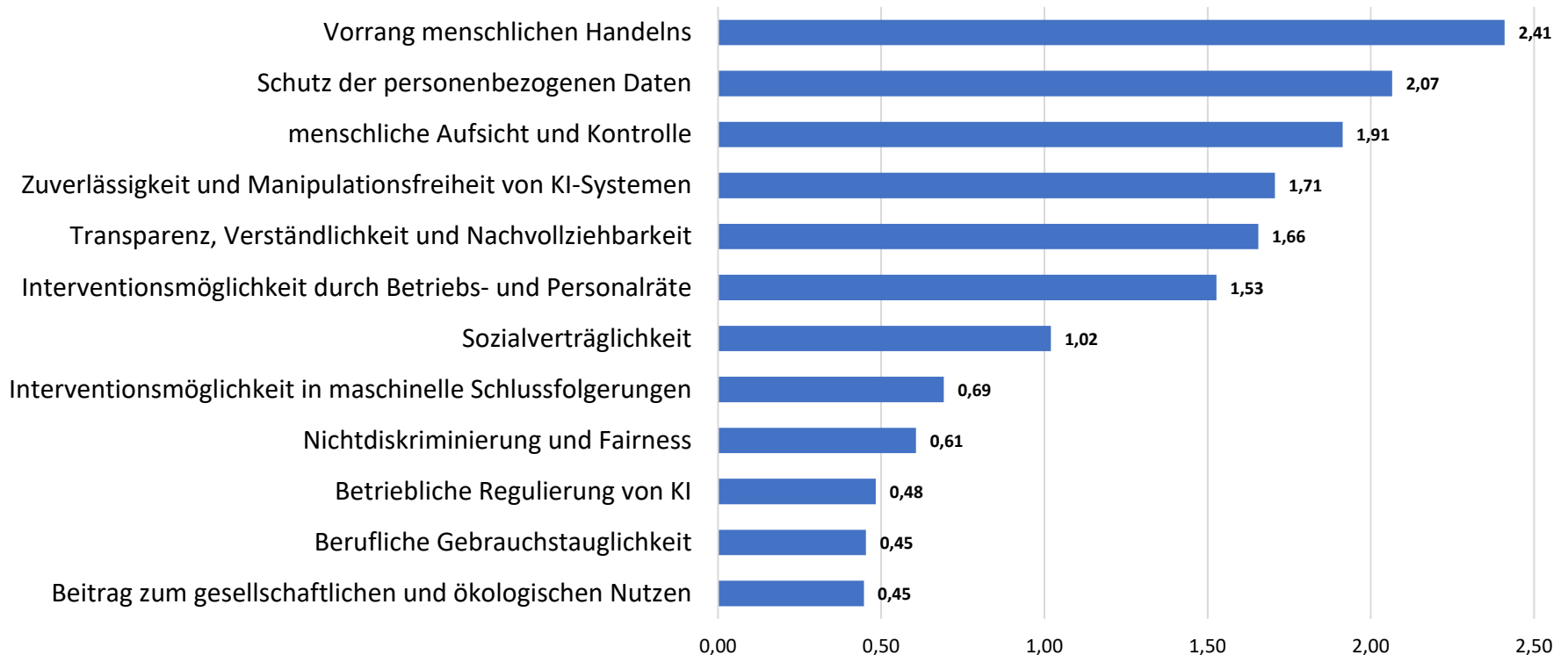


© INPUT Consulting 2021

Quelle : Input- Consulting November 2021

# VERTRAUEN WODURCH? VERTRAUENSFAKTOREN AUS DER SICHT VON BETRIEBS- UND PERSONALRÄTEN

Frage: Was sind aus Ihrer Sicht die wesentlichen Faktoren, an denen Sie das Vertrauen in KI-Technologie festmachen?

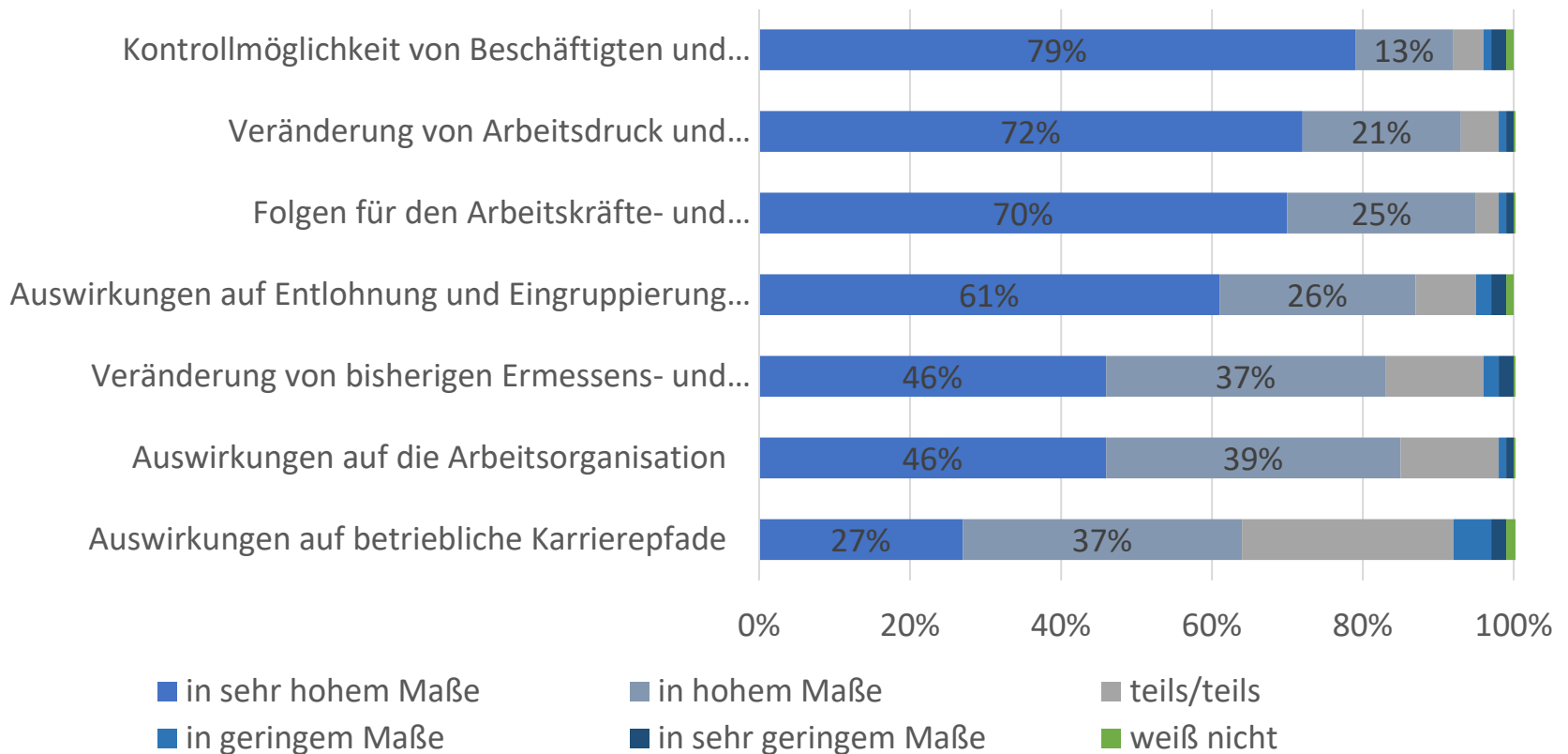


© INPUT Consulting 2021

Quelle : Input- Consulting November 2021

# VERTRAUEN WOMIT ? PROZESSGESTALTUNG: ANFORDERUNGEN AN EINE FOLGENABSCHÄTZUNG

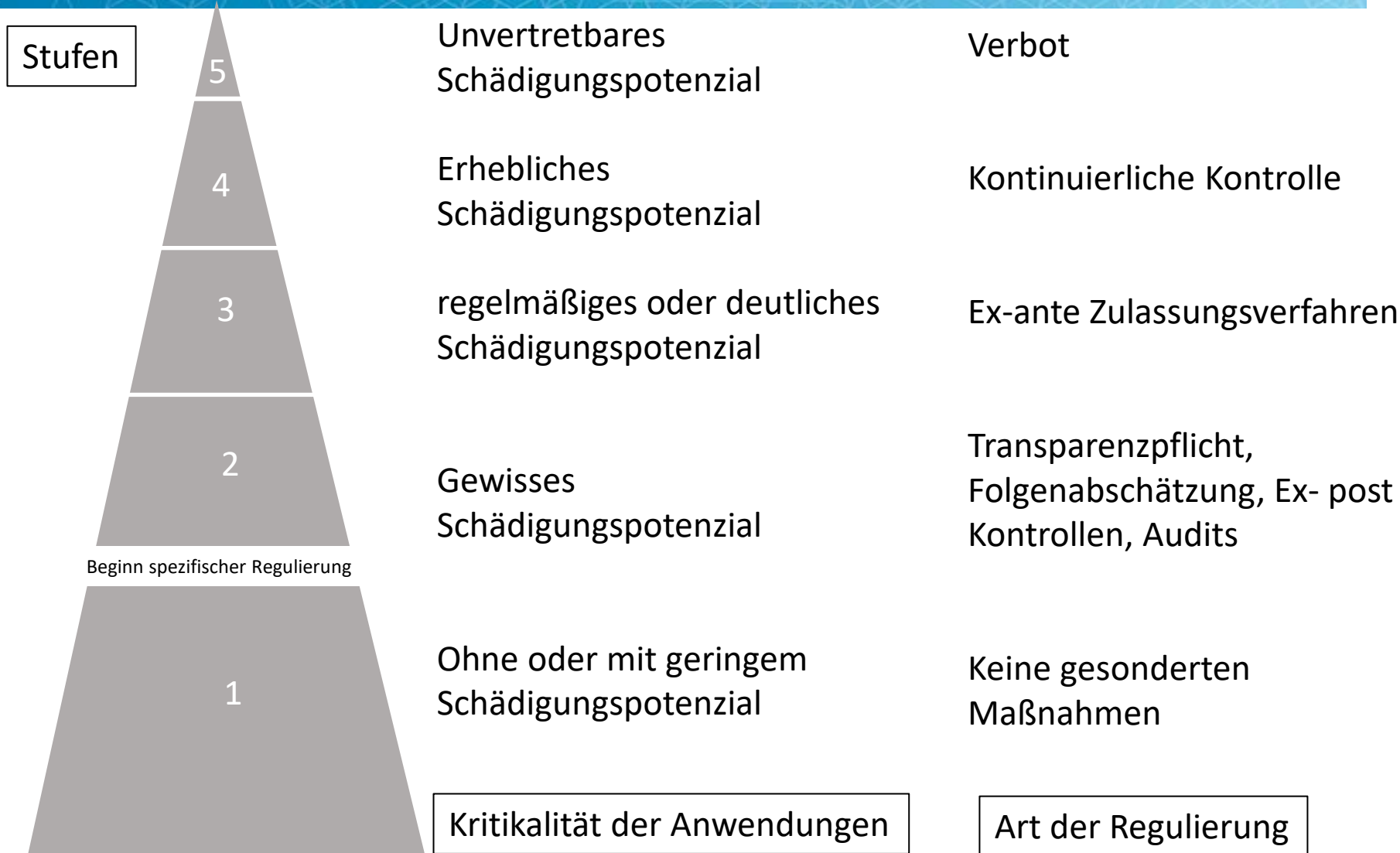
Frage: Welche Aspekte sollten aus Ihrer Sicht bei einer Folgenabschätzung von KI-Systemen in welchem Maße berücksichtigt werden?



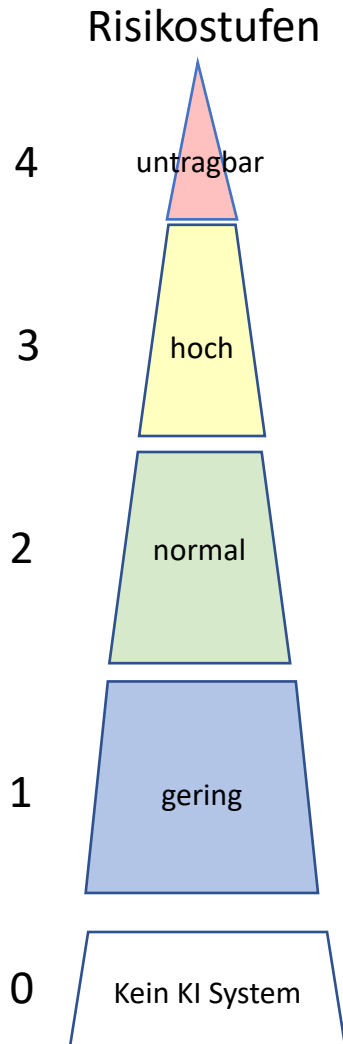
© INPUT Consulting 2021

Quelle : Input- Consulting November 2021

# EMPFEHLUNGEN DER DATENETHIK-KOMMISSION

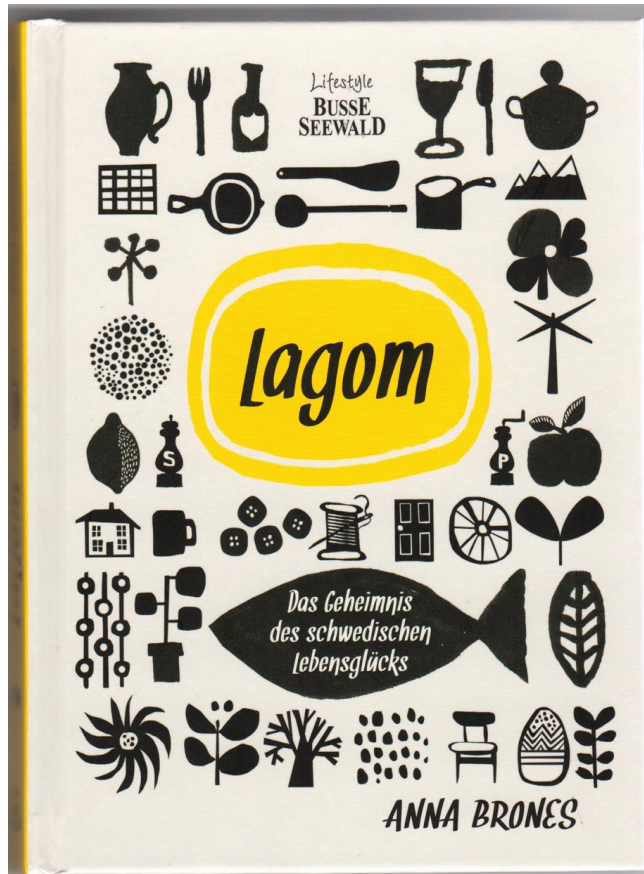


# VERTRAUENSSTIFTENDE PROZESSGESTALTUNG: STRATEGISCHE FILTER FÜR BETRIEBLICHE KRITIKALITÄTSEINSTUFUNG



- Grad maschineller Autonomie
- Schädigungspotenzial
- Grundrechtsrelevanz
- Datenarten
- Zwecke und Ziele
- Erklärbarkeit, Kontrollierbarkeit, Revisions-sicherheit
- Robustheit gegen Manipulation
- Anzahl der Betroffenen
- Funktionsbasis
- Rückholbarkeit

# KULTURELLER VERTRAUENSANSATZ: LAGOM: NICHT ZU VIEL, NICHT ZU WENIG, GENAU RICHTIG



Lagom:

„Für das schwedische Wort lagom gibt es keine genaue Entsprechung.

Lagom kann als Adverb oder als Adjektiv verwendet werden und bedeutet nach dem schwedischen Bedeutungswörterbuch Norsteds Ordbok je nach Kontext genau richtig, ausreichend, in Maßen, gemäßigt, adäquat, passend, geeignet, angemessen oder moderat.

Es wird auch als Substantiv benutzt und beschreibt dann ein Lebensmotto, eine Grundeinstellung oder ein Motto, das sich am guten Mittelmaß orientiert“

KI Lagom



# VERTRAUENSFAKTOREN ALS ORDNUNGSELEMENTE PRIORISIERUNG UND DIALOG ALS ERFOLGSFAKTOR

- Nützlichkeit im Leistungsprozess
- Kontrollierbarkeit
- Rechtskonformität
- Regelkonformität
- Schutz der Persönlichkeitsrechte
- Ergonomie
- Sozialverträglichkeit
- Ethikkonformität
- gute Arbeit
- Robustheit
- Risikoangemessenheit
- Nachhaltigkeit

Wie wichtig ist es für Sie persönlich, dass ... den einzelnen Faktoren im  
Einführungsprozess Rechnung getragen wird ?

Sehr  
wichtig  
1

2

3

4

Nicht  
wichtig  
5

27 Normquellen  
12 Vertrauensfaktoren  
94 Prüfindikatoren  
10 Einführungsphasen  
732 Prüffragen



# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

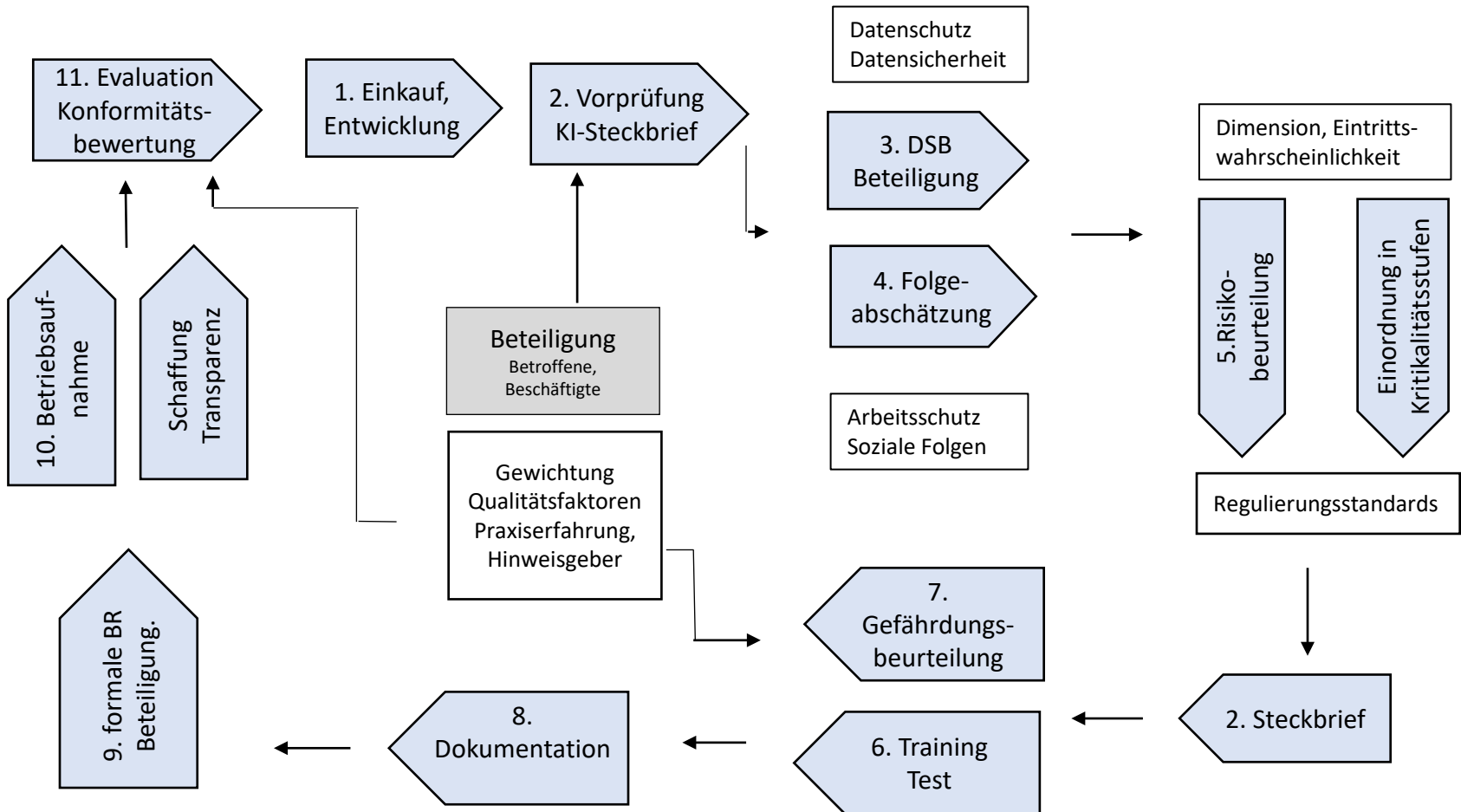
Vertrauensfaktoren	Qualitätsindikatoren		Relevanz
<b>A. Rechtskonformität</b>	A.1	Rechtsgrundlage	A, LI
	A.2	Diskriminierungs- und Verzerrungsfreiheit	A, LI
	A.3	Würde	A
	A.4	Persönlichkeitsentfaltung	B
	A.5	körperliches, geistiges und soziales Wohlergehen	A, LI
	A.6	Recht , Billigkeit und gute Sitten	B
	A.7	Verantwortung	A, LI

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

## C. Nützlichkeit

<b>C.1</b>	Erforderlichkeit	A
<b>C.2</b>	Aufgabenangemessenheit	B
<b>C.3</b>	Zweckdienlichkeit	B
<b>C.4</b>	Gebrauchstauglichkeit	A, LI
<b>C.5</b>	Verfügbarkeit	B
<b>C.6</b>	Belastbarkeit	B
<b>C.7</b>	Effektivität	B
<b>C.8</b>	Effizienz	B
<b>C.9</b>	Wertschöpfungsbeitrag	A, LI
<b>C.10</b>	Innovationsbeitrag	B
<b>C.11</b>	Agilität und Flexibilität	C
<b>C.12</b>	Kundenorientierung	
<b>C.13</b>	Nutzung von Vorerfahrungen	C
<b>C.14</b>	Genauigkeit	B

# VERTAUENSSTIFTENDER EINFÜHRUNGSPROZESS:



# ENQUETEKOMMISSION - VERTRAUENSBIILDUNG: NEUE AUFGABEN FÜR BETRIEBLICHE NORMSETZUNGS AKTEURE :

## Management:

- Schaffung von Transparenz, insbesondere über Zielsetzungen, Datennutzung, Verantwortlichkeiten und Revisionsinstrumente
- Definition von Grenzen, Revisionsmöglichkeiten und verantwortlichen Stellen für maschinelle Schlussfolgerungen
- Entwicklung ethischer Normen
- Einordnung betrieblicher KI-Systeme in Kritikalitätsstufen, Bildung von Nützlichkeitsprofilen
- Durchführung von Tests, Folgeabschätzungen und Gefährdungsanalysen
- Schaffung mitbestimmter Normen
- Ausrichtung des Beschaffungs-, Qualitäts- und Personalmanagements auf betriebliche Normen
- Qualitative Personalplanung und –entwicklung
- Mitwirkung an überbetrieblichem Monitoring und Benchmarking

## Aufsichtsräte:

- Einfordern ethischer Normen
- Beaufsichtigung der Rechtskonformität
- Qualitative Personalplanung

# OPTIONEN FÜR DIALOGE UM VERTRAUENSPRINZIPIEN ZU ENTWICKELN

- Qualitäts- und Vertrauensfaktoren
- Phasen der Qualitätsprüfung und Beteiligung (Vorprüfung, Folgeabschätzung, Risikobewertung, Training, Gefährdungsbeurteilung und Evaluation)
- Erarbeitung von Checklisten für standardisierte Prüfprozesse
- Entwicklung von Ethikstandards
- Klassifizierung nach gemeinsamen Risiko- und Nützlichkeitsklassen
- Standards für personalrelevante Entscheidungen
- Primat menschlicher Autonomie
- Transparenz, Beweisverwertungsverbot, Qualifizierung

## Rahmen für Experimente :

- Expertenkreis mit Auftrag zur Entwicklung von Prüfstandards
- Strategische Filter zur Kritikalitätseinstufung
- Regulierungsstandards für verschiedene Kritikalitätsstufen
- Negativliste für nicht angestrebten KI-Einsatz
- Orte der Datenhaltung und des Datenzugriffs
- Beteiligungsformen für verschiedene Arten von Systemen
- Standards zu Speicherfristen
- Standards zu technisch-organisatorischen Maßnahmen
- Verfahren zur Folgeabschätzung

# DAS KI-LAGOM- KONZEPT DIENT DER VERTRAUENSBIILDUNG

- Verständlich für Betroffene und Experten
- Strukturierungshilfe in der Vielfalt der Aspekte
- Fördert Dialoge
- stützt Eigenständigkeit der Entscheidungsfindung
- Erleichtert Priorisierung
- Verträglich mit unterschiedlichen Anwendungsszenarien
- Lässt Vertrauen entstehen
- Erleichtert Prinzipienfindung in Ethik, Kultur und Vorgehen
- Spart Zeit durch die Vermeidung von Missverständnissen
- Hilft Standards zu bilden und Gestaltungsarbeit zu erleichtern
- Fördert iteratives Vorgehen
- Baut auf bereits legitimierten Vorgaben, Normen und Expertenansforderungen auf
- Erlaubt Vergleiche

# PRAXISHANDBUCH KI-LAGOM



- Warum:** Wirkungen  
Besonderheiten  
Definitionen  
Funktionsarten
- Wofür:** Ethische Maßstäbe  
Anspruchslagen  
Leitideen  
Qualitätsmodell
- Wie:** Vertrauensfaktoren  
Vorgehensschritte  
Beteiligungsvorschläge
- Womit:** Rechtsgrundlagen  
Checklisten  
Prüffragen  
Gute Beispiele



**BACKUP**

# VORPRÄGUNGEN UND VORBELASTUNGEN VERLANGEN NACH EINEM STRUKTURMODELL

- Bilder in Kopf
- 50 Jahre IT-Gestaltungserfahrungen
- Vermeintlich universelle Antworten: Transparenz, Ethik, Verzerrungsvermeidung
- Diskrepanz zwischen Kunden und Beschäftigteneinstellungen (Risiko und Nutzenorientierung)
- Kaum Referenzbeispiele in Gestaltungsverabredungen
- Praktische Hilfen sind rar
- Komplexität und tendenzielle Undurchschaubarkeit
- Aspektvielfalt
- Dominanz anderer Themen
- Nicht nur „informationelle Selbstbestimmung“ ist als Persönlichkeitsrecht relevant

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>B. Ethik</b>	B.1	Kennzeichnung von KI-Systemen	A
	B.2	Primat menschlicher Entscheidungen	A, LI
	B.3	Transparenz und Nachvollziehbarkeit	A, LI
	B.4	Interventionsmöglichkeiten	A
	B.5	Fairness und Diversität	B

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>D. Kontrollierbarkeit</b>	<b>D.1</b>	Dokumentation	B
	<b>D.2</b>	Präzision und Zeitbindung der Ziele und Zwecke	A, LI
	<b>D.3</b>	Integrität der Daten	B
	<b>D.4</b>	Erklärbarkeit	B
	<b>D.5</b>	Steuerbarkeit	A, LI
	<b>D.6</b>	Integration von Sicherheitsmaßnahmen	A
	<b>D.7</b>	Evaluationskonzept	B
	<b>D.8</b>	Verhinderung Fehlanwendung	C
	<b>D.9</b>	Notabschaltung und Alternativkonzept	A
	<b>D.10</b>	Berechtigungskonzept	A, LI

# ZUORDNUNG DER VERTRAUENSAKTOREN ZU PRÜFINDIKATOREN

<b>E. Regelkonformität</b>	E.1	Verordnungen	A, LI
	E.2	DIN, ISO, EN-Normen	A
	E.3	interne Leitlinien und Regeln von Aufsichtsbehörden	A
	E.4	Zertifizierungen	A
	E.5	Tarifvorgaben	A
	E.6	Betriebsvereinbarungen	A
	E.7	Unfallverhütungsvorschriften	A
	E.6	Intensität und Rechtzeitigkeit der Beteiligung	A, LI

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>F. Persönlich- keitsschutz</b>	F.1	Schutz personenbezogener Daten	A, LI
	F.2	Autonomie der Betroffenen	B
	F.3	Privatsphäre im Denken und Fühlen	A
	F.4	Reglementierung der Überwachungspotenziale	A
	F.5	Unterlassung von Profiling und automatisierten Entscheidungen	A, LI
	F.6	Beschränkung personalwirksamer Schlussfolgerungen	A
	F.7	Wahrung der Betroffenenrechte	B

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

## G1 Ergonomie

G.1	Unversehrtheit	A, LI
G.2	Individualisierbarkeit	B
G.3	Barrierefreiheit	B
G.4	Arbeitsplatzgestaltung	C
G.5	Güte der Arbeitsmittel	B
G.6	Arbeitsverfahren	A
G.7	Arbeitszeiten	B
G.8	Qualifikation	B
G.9	psychische Belastung	A,LI
G.10	Aktualität des arbeitswissenschaftlichen Erkenntnisstandes	C
G.11	Wirksamkeit der Arbeitsschutzmaßnahmen	B
G.12	Softwareergonomie	B
G.13	Leistungsanforderungen	A,LI

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>H. Sozial- verträglichkeit</b>	H.1	Soziale Folgenabschätzung	A, LI
	H.2	Inklusion	B
	H.3	Beschäftigungsförderlichkeit	A
	H.4	Altersverträglichkeit	C
	H.5	Kontaktförderlichkeit	B
	H.6	Belastungsminimierung	A



# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>I. Gute Arbeit</b>	I.1	Arbeitszeitsouveränität	C
	I.2	Zugang	C
	I.3	Mischarbeit	B
	I.4	Arbeitserleichterung	A, LI
	I.5	Wertstatus	A, LI
	I.6	Personalentwicklung	B
	I.7	Lernförderlichkeit	C
	I.8	Aufgabenklarheit	B
	I.9	Arbeitsklima	B

# ZUORDNUNG DER VERTRAUENSAKTOREN ZU PRÜFINDIKATOREN

<b>J. Robustheit</b>	J.1	Fehlertoleranz	B
	J.2	Missbrauchsschutz	A
	J.3	Erwartungskonformität	A, LI
	J.4	Manipulationsfestigkeit	A
	J.5	Sicherheitsmaßnahmen	A

# ZUORDNUNG DER VERTRAUENSFAKTOREN ZU PRÜFINDIKATOREN

<b>K. Risiko- angemessenheit</b>	K.1	Risikoeinschätzung	B
	K.2	Risikobewertung	A
	K.3	Risikoklassifizierung	A
	K.4	Risikobewältigung	A
	K.5	Einordnung in Kritikalitätsstufen	A, LI

<b>L. Nachhaltigkeit</b>	L.1	nachhaltiger Ressourceneinsatz	B
	L.2	Emissionsminderung und nachhaltiges Energiemanagement	A