

Foliensatz dient ausschließlich der Information für
wissenschaftliche, nicht kommerzielle Zwecke

Gesund mit KI?

Wesentliche Herausforderungen für eine konstruktive KI-Nutzung

Colloquium, 24.06.2025, Gesundheit
Österreich GmbH, Wien (online)



Gängige Narrative zu Künstlicher Intelligenz

Fragwürdiges Meta-Narrativ

0: KI ist besser als der Mensch in ... XY

- KI steigert Effizienz & Produktivität
- ... bringt Wettbewerbsvorteile & Innovation
- ... verbessert die Kundenzufriedenheit
- ... macht keine Fehler und ist zuverlässig
- ... erhöht Prozess-Qualität, senkt den Arbeitsaufwand & schont Ressourcen
- ... entlastet Arbeit, schafft Freiräume und verbessert Arbeitsbedingungen
- ...
- KI bedroht Arbeitsplätze
- ... verstärkt Diskriminierung
- ... gefährdet die Privatsphäre
- ... ist eine Gefahr für die Sicherheit
- ... bedroht Autonomie und Freiheit
- ... verstärkt Ausbeutung & führt zu Entmenschlichung
- ...

**Polarisierung + enorme Erwartungshaltung im Diskurs
erschweren realistischen Umgang
und damit sozialverträgliche Gestaltung & Nutzung der Technologie**

Was ist Künstliche Intelligenz (KI)? – Rolle und Bedeutung

- Ursprünge: KI = "the science and engineering of making intelligent machines" (McCarthy 1955)
- → Erforschen und entwickeln intelligenter Maschinen
- "understanding the nature of intelligent thought" (Buchanan 2005)

- Zeitgemäßer: **Teilgebiet der Informatik** mit dem Ziel, **menschliche Fähigkeiten** (wie logisches Denken, Lernen, Planen etc.) **zu imitieren und automatisiert unterstützen**

- EU-Def.: maschinenbasiertes System, das mit **unterschiedlichem Grad an Autonomie** operieren kann und für explizite oder implizite Ziele Ergebnisse wie **Vorhersagen, Empfehlungen oder Entscheidungen** generieren kann, die das physische oder virtuelle **Umfeld beeinflussen**. (AI Act/OECD)

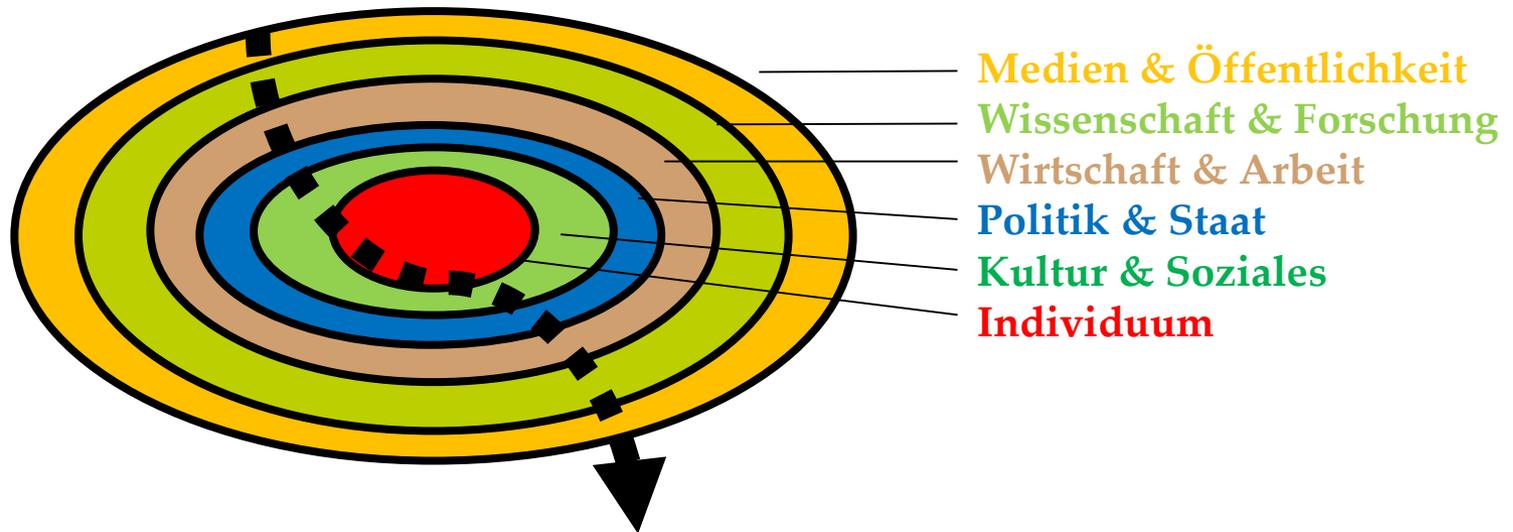
Begriff KI irreführend
→ KI = Einsatz von
mathematisch-statistischen
"Lern"-**Algorithmen**
(Machine Learning) um
Datenmodelle zu trainieren

Frei übersetzt:

- KI=Automatisiertes **System zur Unterstützung** bei **Entscheidungen & Arbeitsprozessen**
- **Mit KI/ML entstehen neue Formen der Automatisierung mit erheblichen Folgen für die Gesellschaft**

Wie verändert KI die Gesellschaft?

Transformationspotenzial auf allen Ebenen



Digitale Technologien/IKT

Datafizierung
Big Data

Algorithmen

KI/Machine
Learning

Digitale Wertschöpfung + Prozessoptimierung
(Effizienz, Effektivität, Produktivität, Schnelligkeit, Skalierbarkeit usw.)
durch Vernetzung und **Automatisierung**

Neue Fragen zur **Formalisierbarkeit** von Arbeitspraktiken

Wie funktioniert KI/ML?

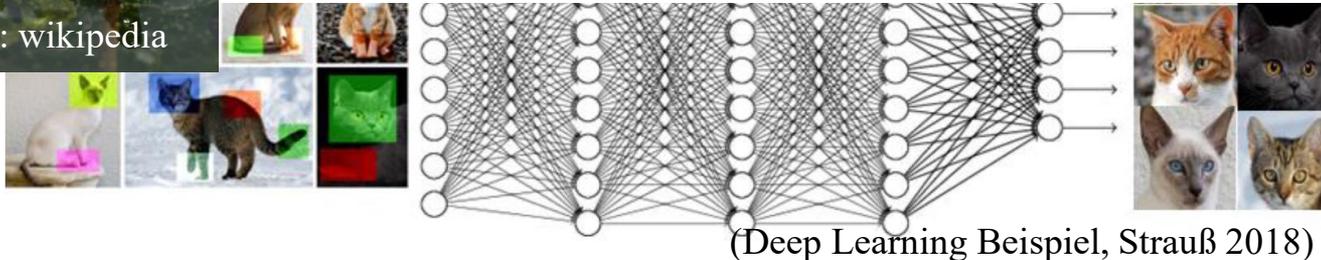
Grundprinzip: Vom Input zum Output



Können Frösche zu Katzen werden?

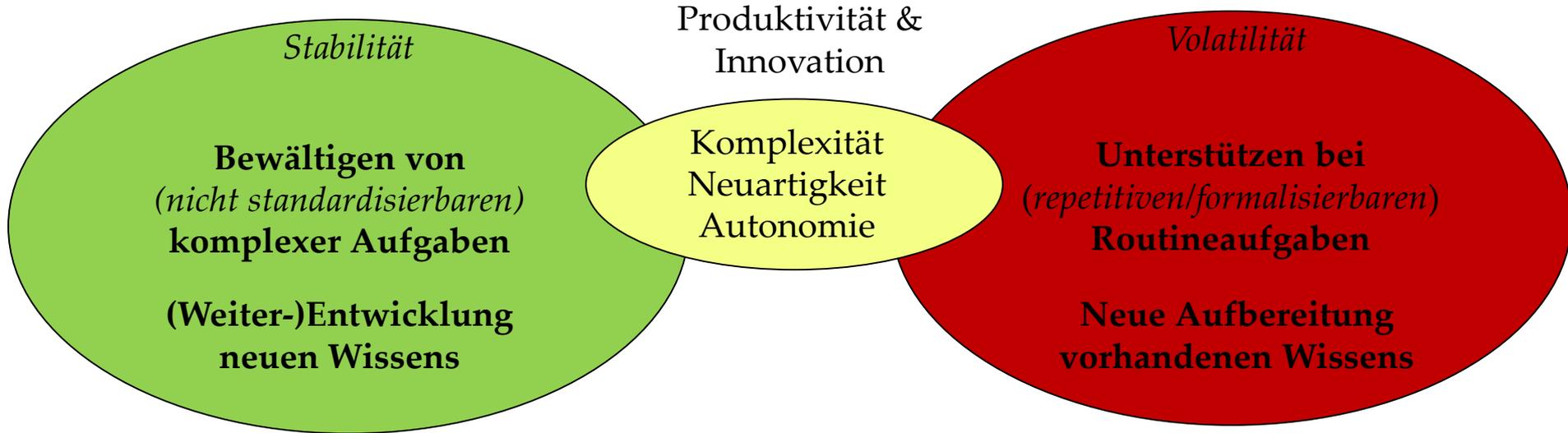


Src: wikipedia



- KI ist gut in: Informations- und **Mustererkennung**, Sprache, Bild, Text, Logik, Daten-Modellierung & **Abbildung von Wissen**
- **Aber:** Qualität von Daten/Input bestimmt Output/Ergebnis und damit die Funktionsfähigkeit eines KI-Systems
- **Qualität, Korrektheit, Kontextualisierung kann KI nicht per se bieten**
→ effizient ≠ korrekt ≠ objektiv

Wissensarbeit und KI: Ein Widerspruch?



Automatisierte Wissensarbeit wäre keine Wissensarbeit mehr
WA ist eine Kombination aus Routine- u. komplexen Aufgaben
Isolierte Sicht darauf kaum praxistauglich (Albrecht/Kellermann 2020)

Wissen ist die Kombination aus (Davenport et al. 1998):

Information, Erfahrung, Kontext, Interpretation & Reflexion

→ **KI kann kein echtes Wissen generieren**

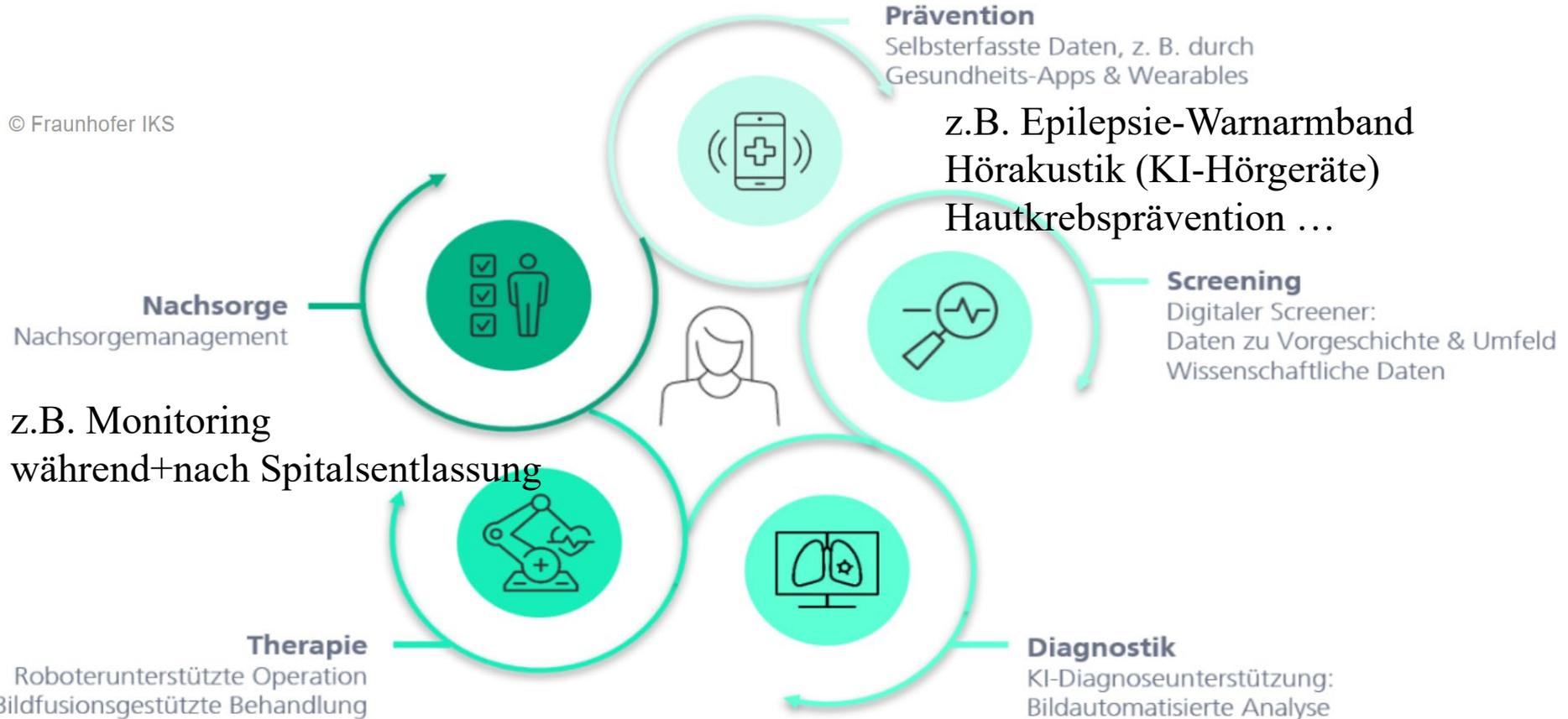
→ Menschliches **Erfahrungswissen und Reflexionsfähigkeit** als besonderer Wert

→ Veränderte Anforderungen im Umgang mit KI in der Wissensarbeit

→ **Critical AI Literacy wird zum Bestandteil moderner Wissensarbeit**

KI in Medizin und Gesundheitswesen

© Fraunhofer IKS



Bildgebung & Messtechnik
Laboruntersuchungen z.B. in
Pathologie, Radiologie, Dermatologie,
Augenheilkunde, Zahntechnik uvm.

KI als Querschnittstechnologie

2 unterschiedliche Hauptformen des KI-Einsatzes

- Virtuell – Aufbereitung und Analyse von Daten
 - Z.B. KI-gestützte Diagnostik, Laboranalysen etc.
- Physisch – Robotik, KI-gestützte physische Behandlung
 - Operationsroboter, Prothesen, Therapieformen etc.
 - Analyse des Patientenzustands, physische Behandlungs-Unterstützung

- → Prozedurale Unterstützung bei Arbeitsabläufen
- → Relevante Zusatzinformationen für Diagnose & Behandlung (z.B. Neben- u. Wechselwirkungen, Krankheitsverläufe, Symptomatik etc.)

Chancen durch KI aus medizinischer Perspektive

Entlastung von Routineaufgaben durch KI-Automatisierung

- Monitoring – KI als „2. Befunder“
- Dokumentation
- Vorsorgeuntersuchungen
- Hervorhebungen (auffällige Fälle)
- ...

Qualitätsverbesserung & Prozessoptimierung

- Informationsmanagement
- Bessere Gerätenutzung
- Höhere chirurgische Genauigkeit
- Objektivierung
- Effizienteres Health Management
- ...

Verbesserte Patientenversorgung

Unterstützung bei Diagnose, Früherkennung & Vorsorge
Überwachung chronischer Krankheiten

Anomalien, „verdeckte“ Symptome, Medikamente-Wechselwirkungen etc.

Risikoreduktion durch Früherkennung u. zeitgerechte Therapieplanung (z.B. MR-Reihung)

Forschung & effizientere Wirkstoffentwicklung

...

Risiken durch KI aus medizinischer Perspektive

Technische Risiken

- Komplexitätssteigerung
- Redundanzen & mangelnde Kontextualisierung
 - Z.b. (teil-)redundante Information ohne Relevanz für Befund, aber Arzt und Patient verunsichern kann.
- Bias & Fehleranfälligkeit (zB. technische Unterschiede, die Diagnose verzerren)

Individuelle & organisatorische Risiken

- Verunsicherung, Autonomie- u. Kontrollverlust
 - Einschränkung der Handlungsfreiheit durch KI-Automatisierung
 - Starke Abhängigkeit („aufgeschmissen, wenn KI nicht funktioniert“)
- Verschlechterung der Patientenversorgung (+KI-Diskriminierung etc.)
- Unklare Verantwortung & Haftungsfragen
- Steigende (soziale und ökonomische) Kosten für Fehlerkorrekturen
- Personelle und finanzieller Mehraufwand bei geringem Nutzen

Hype vs. Realität: Beispiel IBM Watson

2011: IBM computer Watson wins Jeopardy clash (theguardian.com)

→ **Mehrere Pilotprojekte (u.a. in USA, DE, DK) frühzeitig beendet**

(Dt.Krebsforschungszentrum, Houston Hospital, Righshospitalet in Kopenhagen etc.)

Watson weiß nicht weiter (Heise 2019)

Universitätsklinikum Gießen-Marburg **beendet Pilotprojekt 2017 vorzeitig:**

„Performance war inakzeptabel, medizinisches Verständnis bei IBM nicht da“

Onkologen des Rigshospitalet in Kopenhagen lehnten Watson nach einer Weile ab; der

Supercomputer lag mit seinen Empfehlungen für die Krebstherapie zu oft daneben.

Deutsches Krebsforschungszentrum beendete schon 2011 Zusammenarbeit

„Das System ist noch ziemlich unreif. Watson sagt mir zwar, was für eine Behandlung der Patient braucht, aber das ist relativ banal, sozusagen onkologisches Allgemeinwissen“

Prof. Michael Hallek, Direktor Klinik für innere Medizin Uni Köln & Präsident der Deutschen Gesellschaft für Hämatologie und Onkologie (DGHO) (Medizin&Technik 2019)

Watson hilft Radiologen bei Krebsdiagnose (Futurezone 2015)

Dr. Watson hat jetzt Zeit für Sie: IBMs Supercomputer rettet Frau das Leben (Chip 2016)

...

Watson reloaded:

Wie hilfreich sind LLMs/genAI in der Diagnostik?

- ChatGPT **ähnliche Differential-Diagnosen wie Ärzte in Notfallmedizin** (N=30; 87% Übereinstimmung)
Fazit: **kann unterstützend sein, ABER: einige Ergebnisse unplausibel & widersprüchlich** (Berg et al. 2024, Annals of Emergency Medicine)
- randomized clinical trial including 50 physicians, **keine signifikante Verbesserung in Diagnostik durch LLM im Vergleich zu konventionellen Ressourcen** (Goh et al. 2024)
- ChatGPT **nur bei 74 von 150 Fällen richtig**. Studie zeigt **Probleme bei Interpretation von Laborwerten und bildgebenden Verfahren; wichtige Informationen werden übersehen**.
Fazit: **zu ungenau und nicht als Diagnoseinstrument geeignet** (Hadi et al. 2024)
- Tests mit 100 Pädiatrie-Fällen: **Fehlerquote 83%**.
72% völlig falsche Diagnosen; 11% zu breit und ungenau (Barile et al. 2024)
- Neuere Studie Notfallmedizin: **nur marginale Verbesserungen neuerer KI-Modelle**. **Moderat bei Differentialdiagnosen (78%), aber deutlich schlechter bei Hauptdiagnose (60%). Manche Erkrankungen gar nicht erkannt** (z.B. Lungenentzündung) (Wang et al. 2025)

**„One-Catch-All“ Lösungen wie generative KIs
bieten keine Zuverlässigkeit
→ kaum zur Diagnose geeignet**

Spezialisierung längerfristig zielführender

Beispiel: KI in bildgebenden Verfahren zur Diagnose-Unterstützung mit hohem Potenzial

Klassisch: „händische“ Auswertung

von Röntgenbildern (manuelle Sichtung, Analyse, Markierung etc.)

**KI-gestützt: relevante Bereiche
maschinell vormarkiert**

Potenziale: Entlastung monotoner Tätigkeit +
Komplexitätsreduktion durch KI-Support

Vorteil der Bildgebung: visuelle Überprüfbarkeit....

Welche Katze ist fake?

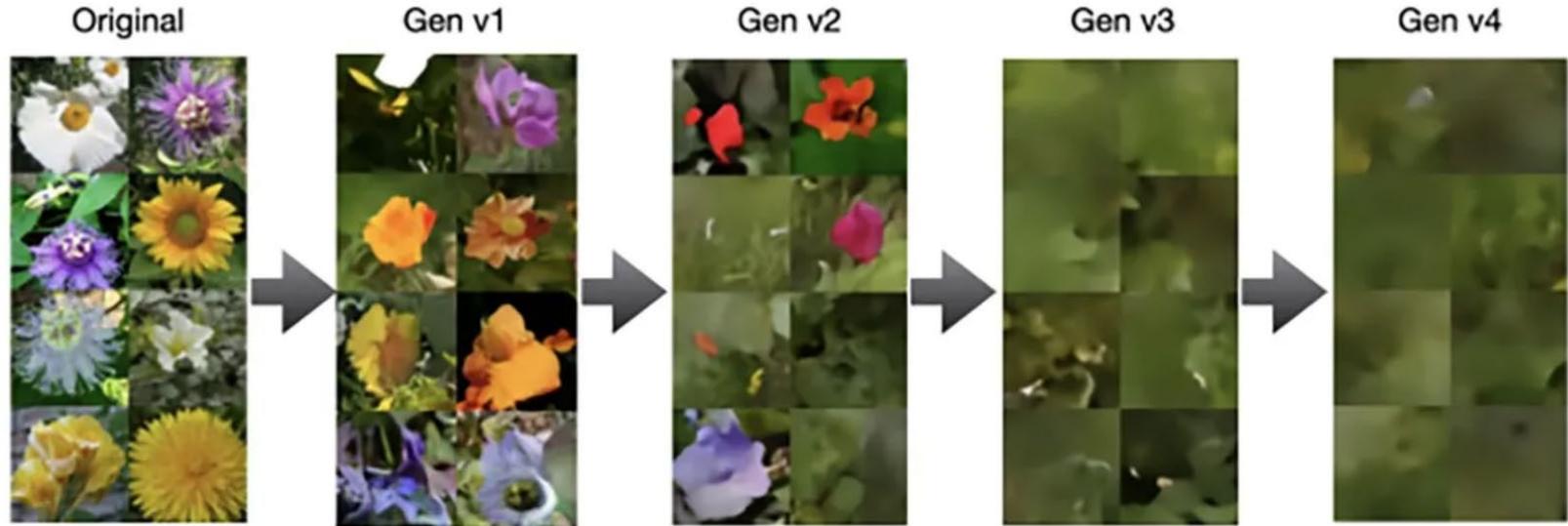


Welche dieser Katzen fängt höchstens digitale Mäuse?

**Stellen Sie sich vor, das wären Röntgenbilder
z.B. der Lunge ...**

<https://data.wdr.de/ddj/deepfake-quiz-erkennen-sie-alle-ki-bilder/>

Model Collapse



An example of what can happen when an AI is trained on AI-generated images. (Credit: G. Martínez et al. / arXiv 2023)

Degenerativer Prozess

Starke Qualitätsverluste, wenn
KI Modell mit rein KI-
generierten (synthetischen)
Daten trainiert wird

Längerfristig Kollapsgefahr
(Shumailov et al. 2024)



Synthetisierung von Daten als komplexe Herausforderung

- KI-Einsatz kann z.B. bestehende Verfahren ergänzen/erweitern (u.a. durch zusätzliche Datenanalysen u. Visualisierungen)
- Z.B. **digital Twinning**: ermöglicht datenbasiertes Patientenabbild
- **Aber: Synthetische Daten können Qualität deutlich verschlechtern**
- Bereits kleine Fehler in Trainingsdaten können fatale Folgen haben „medical LLMs vulnerable to data-poisoning attacks“ (Nature Med. 2025)

→ **Qualitätssicherung von Daten + Prozessen immer wichtiger**

Lehren aus Watson-Fall: starker Bias - Trainingsdaten aus nur 1 US-Klinik

“We had a discussion with [IBM] that they had a very limited view on the international literature, basically, putting too much stress on American studies, and too little stress on big, international, European, and other-part-of-the-world studies,” said Dr. Leif Jensen, who directs the center at Rigshospitalet in Copenhagen that contains the oncology department.

KI zwischen Anspruch und Wirklichkeit

- **Prozessoptimierung** und Effizienzgewinne z.B. durch automatisierte Datenmodellierung, -Visualisierung und -Analysen
 - Verbesserte digitale Wertschöpfung u. Produktion von Inhalten
 - **Arbeitserleichterung** und **Automatisierung** von repetitiven Tätigkeiten
 - **Entscheidungsunterstützung** bei komplexen Aufgaben
- **KI/ML kann helfen, Komplexität zu bewältigen**

→ **Bedeutet aber auch mehr Komplexität**

- Intransparenz, geringe Interpretierbarkeit und höhere Fehleranfälligkeit
- **Abhängigkeit zu KI und Breitenwirkung von Fehlern**
- **Mehr Aufwand & Kosten für Betreuung + Wartung von KI-Systemen**
 - (Datenpflege, Qualitätskontrollen etc.)

→ **KI als Unterstützungs-Tool (Kreativität, Produktivität) begreifen,**
das neue Qualifikationen erfordert
nicht als Ersatz für menschliche Denkleistung

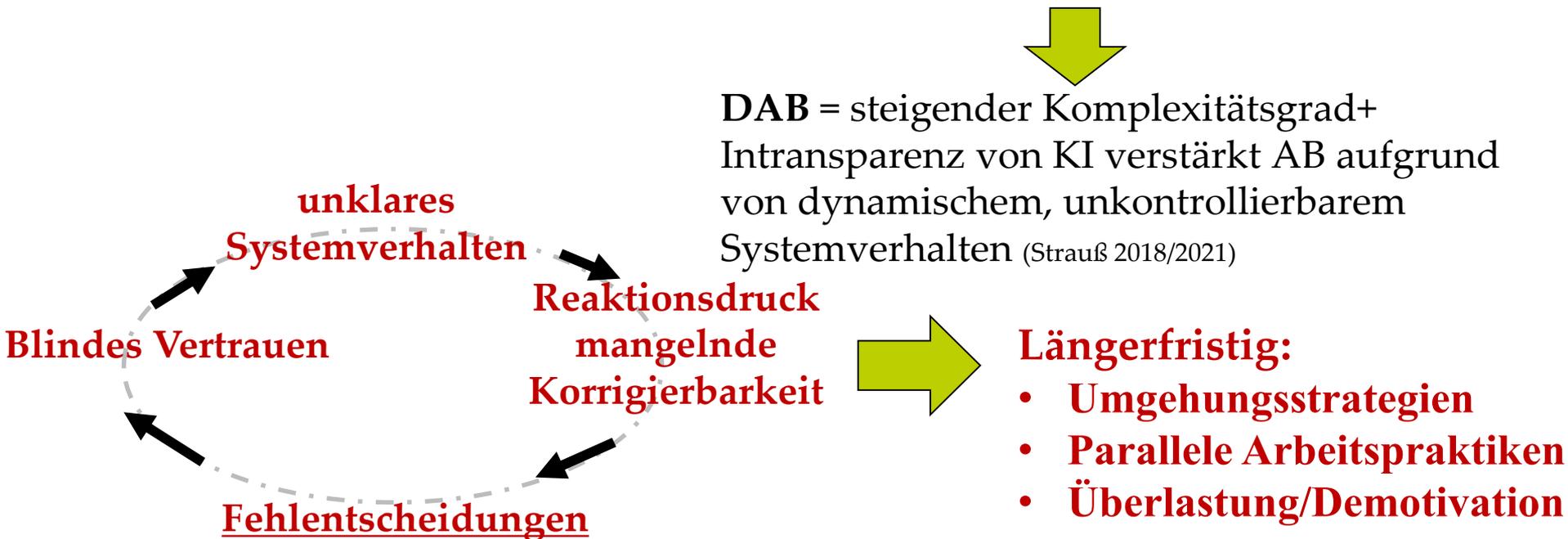
(Deep) Automation Bias: Ein Meta-Risiko der KI-Nutzung

Missverhältnis zw. System-Verhalten und Anwendungspraxis

- AB = Menschliche Neigung, Technologie blind zu vertrauen (cf. Goddard et al. 2014)
→ **Falsche Vorstellungen von Funktionalität & Grenzen automatisierter Technologien**

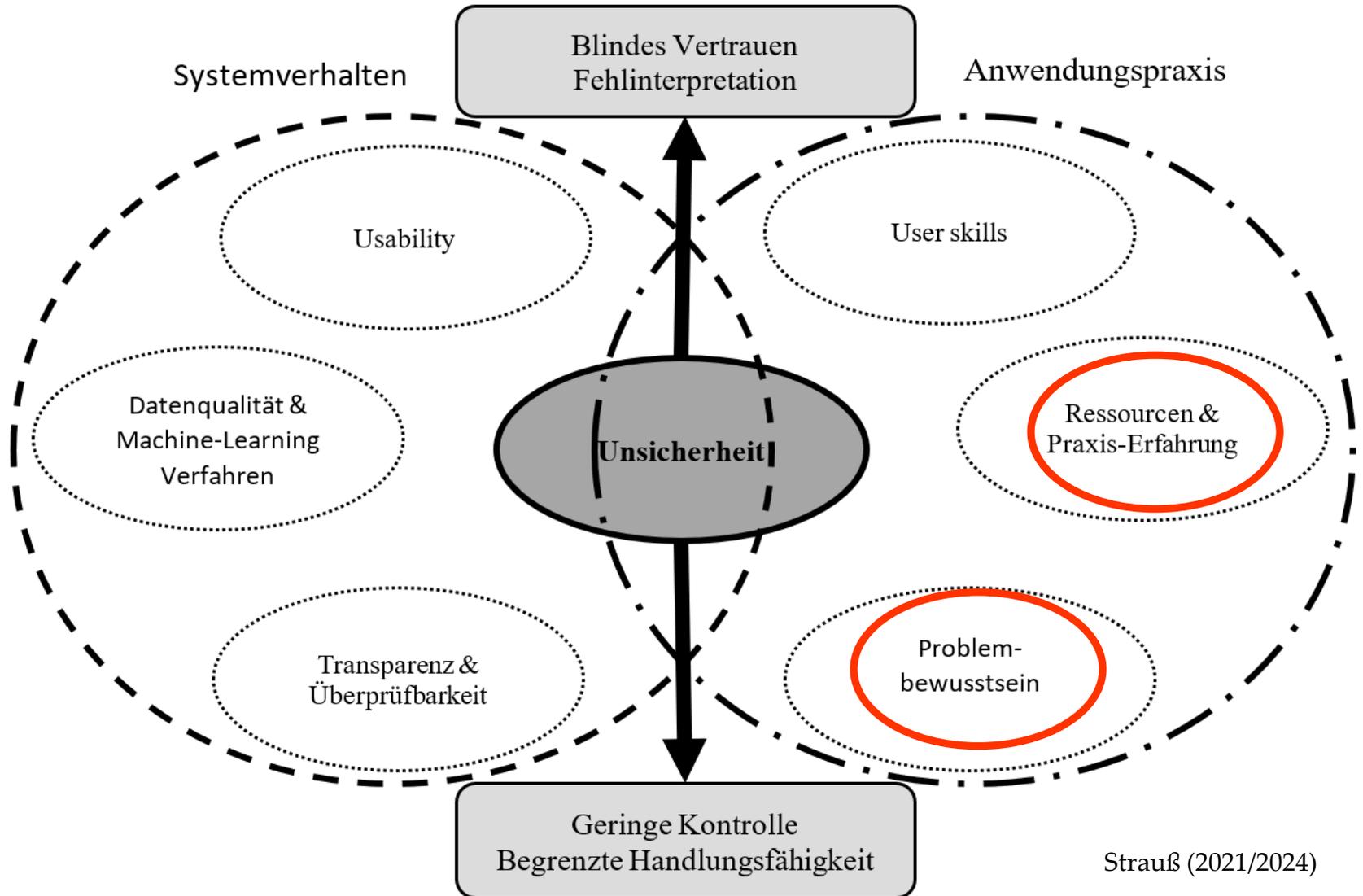
KI/ML erhöht Komplexität & Intransparenz von automatisierten Prozessen

- **Höhere Dynamik, Komplexität & Volatilität**

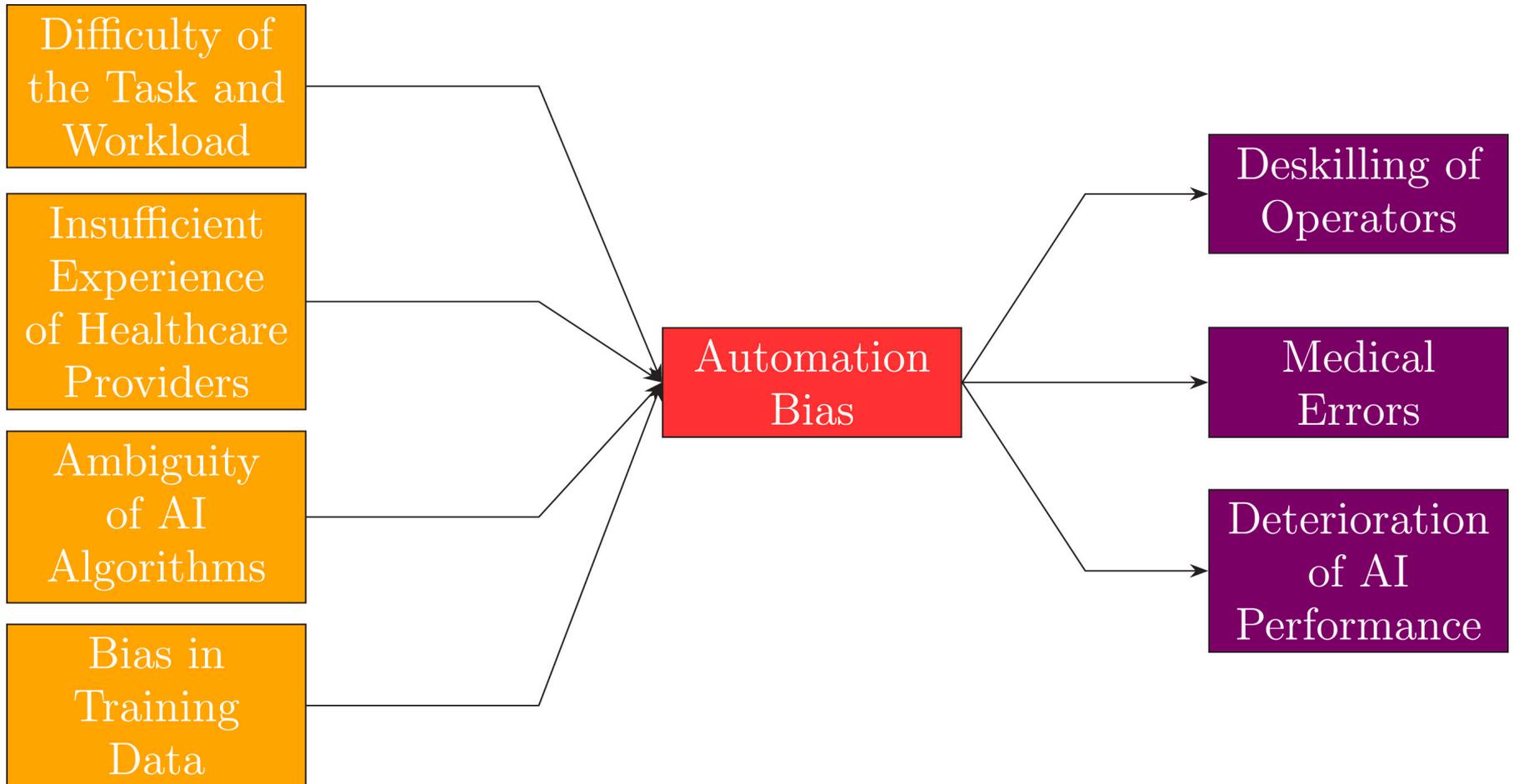


Faktoren, die KI-Automatisierungsrisiken beeinflussen

Missverhältnis zw. System-Verhalten und Anwendungspraxis

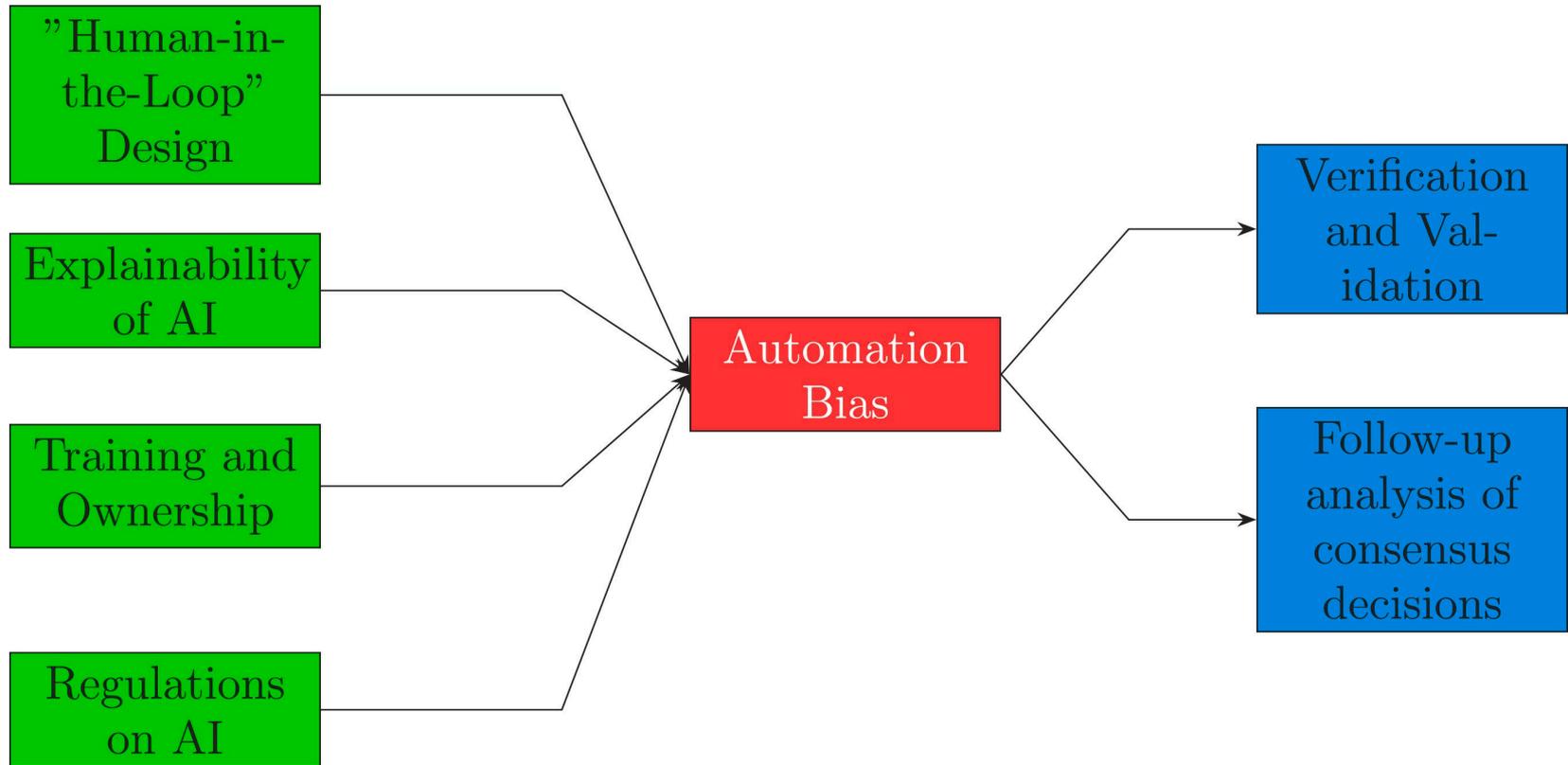


Automation Bias in Healthcare



(Abdelwanis et al. 2024)

Ansätze zur Verringerung von Automation Bias



(Abdelwanis et al. 2024)

Welchen Mehrwert kann KI bringen?

- Vorteile bei Daten-Analyse, -Aufbereitung und **formalisierbaren Tasks**
 - Enormes Potenzial in Medizin zB. bei bildgebender Diagnostik
 - Aber: menschliche Diagnose geht vor KI-gestützter (Tschandl et al. 2020)
- Entscheidend: **Welche Aufgabe soll KI-Einsatz unterstützen?**
- **+ Frage des Mehraufwands** durch KI-Einsatz
- zB. Programmierung: mehr Effizienz möglich, aber ...
 - KI-Einsatz **senkt Qualitätsniveau** von Softwarecodes (Hosbach 2024)
- **Steigende Anforderungen an Qualitätssicherung** (bereichsübergreifend)
 - Unterschätzter Faktor: **ZEIT!**
- **KI-Funktionalität muss mit kontextuellen Anforderungen übereinstimmen**
- **Klar definierte Aufgaben, spezifisches Fachwissen & Verantwortung**

Zentrale Herausforderungen und Spannungsfelder

- Besonderheiten KI-basierter Automation
- Sinnvolle Integration in Arbeitsprozesse
 - Kritischer Faktor Daten-, Anwendungs- und **Prozessqualität**
 - Datenaufbereitung und Datenpflege (manuell, Zusatzaufwand)
- **Mehrwert bei repetitiven Tasks, weniger bei komplexen**
- Kluft zwischen KI als „informed Tool“ in **Expertensystem** vs. „En Passant“ Technologie **ohne klaren Einsatzzweck**
- **Unterschätztes Problem: „Verdeckte“ KI** über Standardsoftware (MS etc.)
- Mehr Aufwand für **Wartung, Qualitätskontrolle & Fehlerkorrektur**
- Akzeptanz und **Akzeptabilität** sind keine Selbstläufer
 - KI kann akzeptiert, aber trotzdem inakzeptabel sein
 - Risiko der **Monotonisierung von Arbeit** durch KI
- Neue Herausforderungen für **Aus-/Fortbildung + Arbeitsbedingungen**

Critical AI Literacy als Voraussetzung für Nutzbarkeit

CAIL=Fähigkeit, die Kernfunktionalität eines KI-Systems + (In-)Kompatibilität in konkreten Anwendungskontexten zu verstehen (Strauß 2021/204)

- **Blindes Vertrauen vermeiden & Handlungsfähigkeit stärken**
- **Problembewusstsein & Basiswissen** zu praktischem Nutzen & Grenzen was kann KI leisten und was nicht?

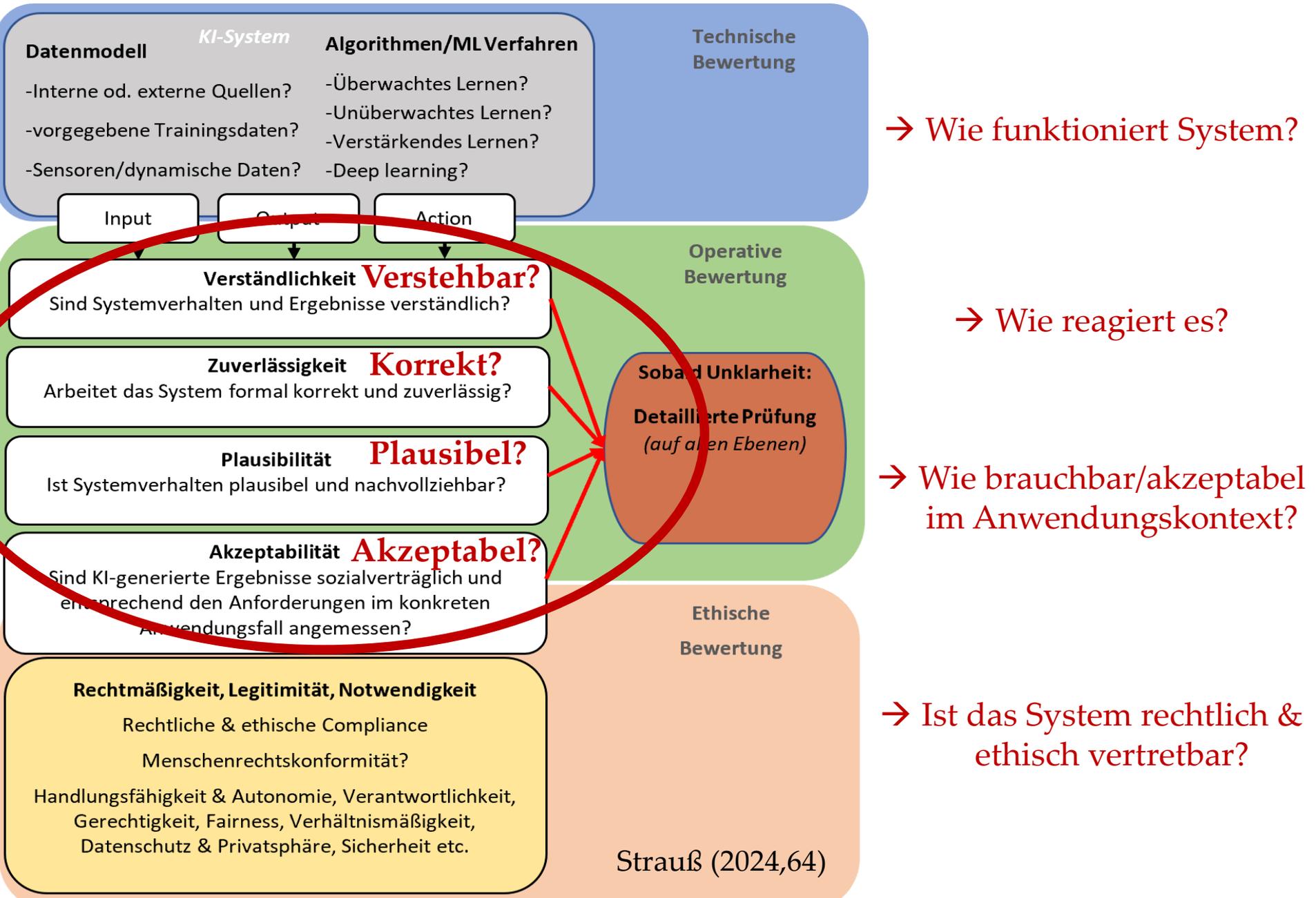
Interpretationsfähigkeit: KI-generierte Inhalte & Ergebnisse

- erkennen & verstehen
→ Aussagekraft und des **situativen Kontexts** (Umstände)
- überprüfen der **Korrektheit** und **Plausibilität**

Bewältigungsfähigkeit: Problemlösungskompetenz im Umgang mit fehlerhaften od. fragwürdigen Ergebnissen

- **beurteilen** von praktischem **Nutzen** und **Akzeptabilität**
→ im konkreten Anwendungsfall brauchbar und akzeptabel oder nicht?
- **verwenden oder verwerfen**

CAIL-Framework zur DAB-Risikominderung



Konstruktiver KI-Einsatz erfordert mehr Differenzierung

Planungsprozess

Wozu KI – Zwecke, Ziele? welches Problem soll es lösen?
Welche Aufgaben soll es unterstützen? Ist das realistisch?

Einführungsprozess

Wie soll KI-System konkret in Arbeitsabläufe **integriert** werden?
Für welche Tätigkeiten/Use Cases? Mit welchen Daten?
Rollen, Verantwortlichkeiten, **Wartung/System/Datenpflege**

Nutzungsprozess

Was ist der praktische Nutzen? Wie im Arbeitsablauf sichtbar?
Bestehendes **Wissen**, Schulungsbedarf, kritische KI-Kompetenz
Problembewusstsein, **Qualitätssicherung**

Adaptionsprozess

Umgang mit Fehlern, Unklarheiten, Anpassungsbedarf
Entscheidungsqualität?
Welche neuen technischen **Abhängigkeiten** entstehen?
Akzeptabilität: Was leistet KI-System zur **Arbeitsverbesserung?**

Zusammenfassung und Ausblick

- KI ist kein Allheilmittel, sondern **beeinflusst Handlungsfähigkeit**
- **KI als Unterstützungs-Tool bei Routineaufgaben begreifen**
- Klare **Kontextualisierung** erleichtert konstruktive Nutzung

Zentrale Frage „**what is the task?**“ (Drucker 99)

Welche Aufgabe soll *wie* mit KI bearbeitet oder unterstützt werden?

Warum und wie wird mit KI automatisiert?
Betriebliche Folgen?



- **Critical AI Literacy als neue Aufgabe der Wissensarbeit**
 - Grundwissen & Problembewusstsein zu KI-basierter Automatisierung
 - Mehr Bewusstsein für Qualitätsanforderungen & **Arbeitsaufwand** durch KI
 - (neue Skills, Qualitätsmanagement, System/Datenpflege etc.) + **ZEIT**

Echter Mehrwert nur mit zweckgebundener & sozialverträglicher Einbettung in Arbeitsabläufe möglich

Vielen Dank für die Aufmerksamkeit!

Dr. Stefan Strauß
Senior Academy Scientist
Institute of Technology Assessment (ITA)
Austrian Academy of Sciences
A-1010 Vienna, Bäckerstraße 13
sstrauss@oeaw.ac.at www.oeaw.ac.at/ita/strauss
Strauß, S. (2024), CAIL Projektbericht
pub.oeaw.ac.at/0xc1aa5572_0x003f893c.pdf
AI.D www.democracy-ai.eu

WWW.OEAW.AC.AT

ÖAW
ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN



PROJEKTBERICHT

Mario A. Pfannstiel *Hrsg.*

Künstliche Intelligenz im Gesundheits- wesen

Entwicklungen, Beispiele
und Perspektiven

Springer Gabler

CAIL – CRITICAL AI LITERACY

KRITISCHE TECHNIKKOMPETENZ FÜR KONSTRUKTIVEN UMGANG
MIT KI-BASIERTER TECHNOLOGIE IN BETRIEBEN

GEFÖRDERT DURCH
Digifonds

