

Wie Menschen entscheiden, wenn Maschinen mitdenken.

– *Über Verantwortung im Zeitalter algorithmischer Assistenz.*

Vier irritierende Thesen zu Automation Bias, der Psychologie der Verantwortungsdiffusion und der Frage, warum „menschliche Aufsicht“ als regulatorisches Konzept an der menschlichen Kognition scheitert.

<p>AUTORIN</p> <p>Silvia Mann-Kundt Systems Thinker · Imago Atelier</p>	<p>EINORDNUNG</p> <p>Whitepaper 04 · Kernthema 04 Herbst 2027 · complexity-organizer.com</p>
--	--

Quellen: Springer AI & Society (2025) · Oxford ISQ (2024) · Oxford JPART (2023) ·
Frontiers in Psychology (2023) · Journal of Business Ethics (2025) · Analyses of Social
Issues & Public Policy (2025) · Scientific Reports (2023) · EU AI Act (2024/1689)

EXECUTIVE SUMMARY

Das Versprechen und sein Riss

Der EU AI Act schreibt menschliche Aufsicht für Hochrisiko-KI-Systeme zwingend vor. Das klingt vernünftig. Es ist aber gleichzeitig eine der ambitionösesten Anforderungen, die ein Regelwerk je an menschliche Kognition gestellt hat. Denn Menschen sind, wie die Forschung zeigt, strukturell schlechte Aufseher über algorithmische Systeme — nicht weil sie unfähig sind, sondern weil das menschliche Gehirn unter Bedingungen von Zeit- und Entscheidungsdruck systematisch und vorhersehbar in Mechanismen verfällt, die Verantwortung erodieren.

Dieses Whitepaper argumentiert: Das Problem mit KI-Entscheidungen ist nicht, dass Menschen Maschinen nicht vertrauen. Es ist, dass Menschen Maschinen häufig zu sehr vertrauen — und dabei die Verantwortung für das Ergebnis systematisch abgeben, ohne es zu merken. Das hat einen Namen: Automation Bias. Und es hat eine psychologische Schwester: die Diffusion von Verantwortung. Beide sind empirisch gut belegt, politisch kaum adressiert, und praktisch in fast jedem KI-Einsatz in Organisationen virulent.

Die vier Thesen dieses Papers sind Irritationen, die gezielt dort ansetzen, wo der KI-Governance-Diskurs zu selbstsicher wird: bei der Annahme, dass menschliche Aufsicht eine Kontrollinstanz ist, statt ein kognitives Problem.

DIE VIER THESEN DIESES PAPERS

- These 1: Automation Bias ist kein Benutzerfehler — er ist eine kognitive Systemreaktion auf Komplexitätsreduktion.
- These 2: „Menschliche Aufsicht“ ist ein Rechtsbegriff, kein Kognitionsbegriff — und diese Verwechslung erzeugt ein falsches Sicherheitsgefühl.
- These 3: Verantwortungsdiffusion im Mensch-Maschine-Team ist keine Ausnahme, sondern der Normalfall — wenn Zuständigkeiten nicht explizit definiert sind.
- These 4: Die Frage „Wer hat entschieden?“ ist in KI-gestützten Prozessen ohne klare Antwort — und das ist gefährlicher als jede schlechte KI-Empfehlung.

Was wir über Menschen und algorithmische Systeme wissen

Automation Bias: Der Befund

Automation Bias (AB) bezeichnet die kognitive Neigung, automatisierten Empfehlungen mehr Vertrauen zu schenken als menschlichen — auch wenn Hinweise auf Fehler vorhanden sind. Eine systematische Review in *AI & Society* (Springer, 2025) analysierte auf Basis des PRISMA-2020-Protokolls 35 peer-reviewed Studien aus den Jahren 2015–2025, die AB in modernen KI-Umgebungen untersuchten. Die Studie umfasst Disziplinen wie kognitive Psychologie, Human Factors Engineering, Human-Computer Interaction und Neurowissenschaften.

AI & Society (Springer, 2025): Exploring automation bias in human–AI collaboration: a review and implications for explainable AI. doi.org/10.1007/s00146-025-02422-7

Zentrale Befunde des Reviews: AB hängt nicht allein von Vertrauen ab. Menschen neigen sowohl zu Übervertrauen (Over-Reliance) als auch zu Überaversion (Algorithm Aversion) — abhängig von Aufgabenkomplexität, persönlicher Expertise und dem Design der KI-Schnittstelle. Besonders relevant: AB ist mit dem Anchoring Effect verknüpft. Wer eine KI-Empfehlung als erste Information erhält, neigt dazu, seine eigene Einschätzung an dieser Empfehlung zu verankern, statt unabhängig zu urteilen.

Eine Studie mit praktisch hoher Relevanz: Horowitz & Kahn (2024, *International Studies Quarterly*, Oxford) befragten in einem präregistrierten Experiment 9.000 Erwachsene in neun Ländern. Ihre Kernthese: Je geringer das tatsächliche KI-Wissen eines Menschen, desto höher die Wahrscheinlichkeit von Automation Bias — bis ein Dunning-Kruger-ähnliches Plateau erreicht ist. Wer wenig weiß, zweifelt am meisten. Wer etwas weiß, vertraut blind. Wer wirklich viel weiß, beginnt wieder zu kalibrieren. Für Organisationen, die KI ohne substantielle KI-Literacy-Programme einsetzen, ist das ein kritischer Befund.

Horowitz, M.C. & Kahn, L. (2024): Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts. *International Studies Quarterly* 68(2), sqae020. doi.org/10.1093/isq/sqae020

AI & SOCIETY 2025 35 peer-reviewed Studien zu AB in KI-Umgebungen	HOROWITZ & KAHN 2024 9.000 Befragte in 9 Ländern, präregistriert	FRONTIERS PSYCH 2023 93 Teilnehmende, Experim. zu AB-Reduktion durch Verantwortungsklarheit	JOURNAL BUS. ETHICS 2025 Gap „Verantwortungsluft“ bei KI-Mensch-Entscheidungen
--	---	--	---

Verantwortungsdiffusion: Das psychologische Muster

Frontiers in Psychology (2023) untersuchte Strategien zur Reduktion von Automation Bias in KI-gestützter Personalauswahl. Kernbefund: Ein zentraler Mechanismus hinter AB ist die Diffusion von Verantwortung. Dieser beschreibt das Phänomen, dass Menschen sich

weniger persönlich verantwortlich fühlen, wenn sie eine Entscheidungsaufgabe mit einem anderen Akteur — ob Mensch oder Maschine — teilen. Bei KI-Systemen tritt dieser Effekt besonders stark auf, weil die Maschine als „objektiv“ und „systemisch“ wahrgenommen wird: Sie scheint zu wissen, was sie tut. Das verleitet dazu, ihre Outputs als Ergebnis und nicht als Empfehlung zu behandeln.

Frontiers in Psychology (2023): Check the box! How to deal with automation bias in AI-based personnel selection. doi.org/10.3389/fpsyg.2023.1118723

Das Journal of Business Ethics (2025) bestimmt das Phänomen konzeptionell als „Verantwortungsluft“ (Responsibility Gap): Obwohl die rechtliche und moralische Verantwortung bei menschlichen Entscheidungsträgern verbleibt, kann die Beteiligung von KI eine psychologische Diffusion von Verantwortung auslösen, die Manager weniger persönlich accountable für potenzielle Schäden fühlen lässt. Das Konzept der „viele Hände“ (Coeckelbergh 2020) beschreibt, wie Verantwortung über mehrere Akteure — Designer, Deployer, Nutzer — verteilt wird, sodass niemand das vollständige Bild hält.

Journal of Business Ethics (2025): The Bright Side of AI in Marketing Decisions: Collaboration with Algorithms Prevents Managers from Violating Ethical Norms. doi.org/10.1007/s10551-025-06083-w

Der Mechanismus des „motivated reasoning“ und der „moralischen Distanzierung“ ist in *Analyses of Social Issues and Public Policy (2025)* systematisch ausgearbeitet. Das Paper argumentiert: KI-Werkzeuge erzeugen psychologische Distanz, die moralische Accountability schwächt. Nutzer können Ungleichheit perpetuieren, während sie ihren Glauben an die eigene Objektivität aufrechterhalten — weil das Ergebnis algorithmisch erscheint. Das nennt die Autorin „AI as moral cover“.

Analyses of Social Issues and Public Policy (2025): AI as moral cover: How algorithmic bias exploits psychological mechanisms to perpetuate social inequality. doi.org/10.1111/asap.70031

Die gefährlichste KI-Eigenschaft ist nicht, dass sie falsch liegen kann. Es ist, dass sie Menschen glauben lässt, nicht selbst entschieden zu haben — obwohl sie es haben.

Was der öffentliche Sektor lehrt

Ein besonders aufschlussreicher Befund stammt aus der Forschung zum öffentlichen Sektor. Eine Studie im *Journal of Public Administration Research and Theory* (Oxford, 2023) untersuchte „Selective Adherence“: Entscheidungsträger folgen algorithmischen Empfehlungen selektiv — sie nehmen sie an, wenn sie ihre vorhandenen Annahmen bestätigen, und ignorieren sie, wenn sie diesen widersprechen. Das Ergebnis ist kein neutraler Automatismus, sondern eine Verstärkung vorhandener Vorannahmen — unter dem Deckmantel algorithmischer Objektivität.

Journal of Public Administration Research and Theory (Oxford, 2023): Human–AI Interactions in Public Sector Decision Making: ‘Automation Bias’ and ‘Selective Adherence’ to Algorithmic Advice. doi.org/10.1093/jpart/muac033

Was Verantwortung, Kontrolle und Vertrauen bedeuten – präzise

Eine der Grundfragen der Rationalität lautet: Wie lauten die Definitionen der zentralen Begriffe? Governance-Dokumente, KI-Richtlinien und Unternehmenspolicies sind voll von Begriffen wie Verantwortung, Kontrolle, Überwachung und Vertrauen. Die wenigsten Dokumente definieren diese Begriffe – und noch weniger prüfen, ob die Beteiligten unter diesen Begriffen dieselbe Sache verstehen.

Verantwortung	im Fehlerfall wissen, was zu tun ist — und die Kompetenz haben, es zu tun. Nicht: formal zuständig sein. Nicht: im Organigramm stehen. Sondern: wissen, handeln können.
Kontrolle	die Folgen einer Entscheidung auf den Schirm kriegen & bewerten können. Nicht: alle Schritte überwachen. Sondern: Ergebnisse wahrnehmen und einordnen können.
Vertrauen	im Sinne von Vereinbarungen handeln; Fremdkontrolle unterlassen und dort, wo Kontrolle vorgefunden wird, durch Selbstkontrolle ersetzen. Nicht: blind glauben. Sondern: begründete Erwartung, die Transaktion reduziert.
Menschl. Aufsicht	die Kapazität, eine KI-Empfehlung zu beurteilen, ihr zuzustimmen oder sie abzulehnen — basierend auf eigenem Urteilsvermögen, das auf Domänenwissen beruht. Nicht: lesender Blick auf ein Dashboard.

Diese Definitionen sind bewusst operativ formuliert — nach dem Prinzip: Eine gute Definition nennt den Geltungsbereich, den nächsthöheren Oberbegriff und grenzt gegen Nachbarbegriffe ab, ohne Zirkelschlüsse. Sie sind auch bewusst anspruchsvoller als das, was in den meisten Unternehmensrichtlinien zu lesen ist.

Warum das wichtig ist: Wenn Verantwortung als formale Zuständigkeit definiert wird, kann man sie dokumentieren. Wenn Verantwortung aber im Fehlerfall tatsächlich gehandelt werden muss, reicht Dokumentation nicht. Und der EU AI Act verlangt das Letztere — auch wenn er es nicht so präzise formuliert.

Systems-Thinking-Perspektive

Entscheidungen sind im systemtheoretischen Sinne Kommunikationen, die Alternativen schließen. Wenn KI eine Empfehlung gibt und ein Mensch ihr folgt, hat der Mensch entschieden — nämlich: keine eigene Prüfung vorzunehmen. Diese Entscheidung ist genauso folgenreich wie die inhaltliche Entscheidung selbst. Governance muss beide adressieren.

Was die Governance-Debatte übersieht

Die folgenden vier Thesen sind produktive Irritationen. Sie beruhen auf empirischer Forschung — und sie sind absichtlich so formuliert, dass sie die Selbstverständlichkeiten des KI-Governance-Diskurses stören. Widerspruch ist ausdrücklich erwünscht.

THESE 01 **Automation Bias ist kein Benutzerfehler — er ist eine kognitive Systemreaktion auf Komplexitätsreduktion.**

Der gängige Umgang mit Automation Bias in Unternehmen: Schulungen. „Nutz die KI kritisch“. „Prüfe die Empfehlung“. Das ist naiv, weil es Automation Bias als individuelles Versagen behandelt. Die Forschung zeigt aber eindeutig: AB ist eine systemische Reaktion. Das menschliche Gehirn ist für Komplexitätsreduktion optimiert. Wenn ein System eine plausible Antwort liefert, reduziert das kognitive Last — und das Gehirn ergreift diese Entlastung. Das ist kein Bug im Nutzer. Das ist ein Feature der menschlichen Kognition, die unter Zeitdruck und hoher Entscheidungslast immer heuristisch arbeitet. Das bedeutet: Wer Automation Bias durch Schulung eliminieren will, kämpft gegen die Kognitionsarchitektur des Menschen. Wer ihn durch Systemdesign reduzieren will, hat eine realistische Chance.

AI & Society (Springer, 2025): Exploring automation bias in human–AI collaboration / Frontiers in Psychology (2023): Check the box! / Horowitz & Kahn (2024): Bending the Automation Bias Curve.

THESE 02 **„Menschliche Aufsicht“ ist ein Rechtsbegriff, kein Kognitionsbegriff — und diese Verwechslung erzeugt falsches Sicherheitsgefühl.**

Art. 14 EU AI Act verlangt für Hochrisiko-KI-Systeme, dass Deployer sicherstellen, dass natürliche Personen eine effektive Aufsicht ausführen können. Das ist als Rechtspflicht formuliert. Aber was eine „effektive Aufsicht“ kognitiv voraussetzt, ist im Gesetz nicht definiert — und in den meisten Unternehmen nicht operationalisiert. Effektive Aufsicht setzt voraus: Domänenwissen, das eine Plausibilitätsprüfung der KI-Empfehlung erlaubt. Zeit und kognitive Kapazität für diese Prüfung. Explizite Verantwortung für die Entscheidung. Und die Bereitschaft, die KI-Empfehlung abzulehnen. Keiner dieser Faktoren ist durch eine Compliance-Prüfung messbar. Was Compliance dokumentiert, ist die formale Existenz einer Aufsichtsperson. Was sie nicht dokumentiert: ob diese Person tatsächlich in der Lage ist, die Empfehlung zu beurteilen.

Verordnung (EU) 2024/1689 (EU AI Act), Art. 14 / AI & Society (Springer, 2025) / Journal of Public Administration Research (Oxford, 2023).

THESE 03 Verantwortungsdiffusion im Mensch-Maschine-Team ist keine Ausnahme, sondern der Normalfall — wenn Zuständigkeiten nicht explizit definiert sind.

Forschung zum „many hands problem“ (Coeckelbergh, 2020) und zur Verantwortungsdiffusion bei menschlich-automatisierten Entscheidungen (Skitka et al., 2000) zeigt konsistent: Menschen fühlen sich weniger verantwortlich, wenn ein anderer Akteur beteiligt ist — auch wenn dieser Akteur eine Maschine ist. Das Journal of Business Ethics (2025) bestätigt: Die Beteiligung von KI induziert psychologische Diffusion von Verantwortung, die Manager weniger persönlich accountable für potenzielle Schäden fühlen lässt. Diese Diffusion ist nicht böser Wille. Sie ist die automatische Konsequenz geteilter Entscheidungsaufgaben. Die Lösung ist nicht Motivation, sondern Strukturklarheit: Wer entscheidet was? Wer trägt die Folgen im Fehlerfall? Und weiß diese Person das — bevor das KI-System aktiviert wird?

Journal of Business Ethics (2025) / Frontiers in Psychology (2023) / Analyses of Social Issues & Public Policy (2025): AI as moral cover.

THESE 04 Die Frage „Wer hat entschieden?“ ist in KI-gestützten Prozessen ohne klare Antwort — und das ist gefährlicher als jede schlechte KI-Empfehlung.

Wenn ein Personalverantwortlicher eine Kandidatin ablehnt, weil der KI-Algorithmus eine niedrige Eignung signalisiert hat, wer hat dann entschieden? Juristisch: der Mensch. Psychologisch: zum Teil die Maschine. Faktisch: die Frage ist unbeantwortet. Genau diese Unklarheit ist das Kernproblem. Schlechte KI-Empfehlungen können korrigiert werden, sobald sie identifiziert sind. Ungeklärte Verantwortlichkeit hingegen erzeugt systemische Fehler, die nicht erkannt werden — weil niemand sich verantwortlich fühlt, sie zu sehen. Die Preprints-Studie (2025) beschreibt diesen Zustand als „Attributionslücke“: Decision Support Systems redistribuieren Autonomie und Verantwortlichkeit, ohne dass Institutionen ihre Accountability-Modelle entsprechend angepasst hätten.

The Impact of AI-Based Decision Support Systems on Human Agency in Governance (Preprints, Nov. 2025) / Journal of Public Administration Research (Oxford, 2023).

Verantwortung greift immer da, wo Fehler passieren. Wenn niemand im Fehlerfall weiß, was zu tun ist — oder wer verantwortlich ist — ist die Governance nicht nur unvollständig. Sie existiert nicht.

Von der Aufsichtspflicht zur Aufsichtskompetenz

Die Konsequenz aus den vier Thesen ist klar: KI-Governance, die sich auf die formale Benennung von Aufsichtspersonen beschränkt, produziert ein Sicherheitsgefühl ohne Sicherheit. Was stattdessen nötig ist, sind drei operative Verschiebungen.

Verschiebung 1: Von Dokumentation zu Entscheidungsklarheit

Die wichtigste Frage vor dem Einsatz eines KI-Systems ist nicht: „Haben wir alle Felder der Compliance-Dokumentation ausgefüllt?“ Es ist: „Wer kann die KI-Empfehlung dieses Systems mit eigenem Urteilsvermögen beurteilen, und wer trägt im Fehlerfall die Verantwortung — und weiß diese Person das?“ Diese Frage muss für jedes KI-System, das Entscheidungen unterstützt, explizit und namentlich beantwortet sein.

Das ist anspruchsvoller als eine Risikoklassifikation. Es verlangt, dass Organisationen prüfen, ob ihre „Aufsichtspersonen“ tatsächlich über die Domänenexpertise verfügen, die nötig ist, um die Empfehlung zu beurteilen. Wenn nicht, ist die Aufsicht formal vorhanden, aber funktional leer.

Verschiebung 2: Systemdesign gegen Automation Bias

Die Forschung zeigt, dass Automation Bias durch Systemdesign reduziert werden kann, aber nicht durch Schulungen allein. *Frontiers in Psychology* (2023) identifiziert zwei wirksame Designmaßnahmen: explizite Hinweise auf die Möglichkeit von Systemfehlern, und eine Formulierung, die die Verantwortung des Entscheidenden in den Vordergrund rückt. Auf der Ebene des Elaboration Likelihood Model ausgedrückt: Wer systematisch denken soll, braucht Bedingungen, die systematisches Denken fördern — nicht heuristisches.

Konkret: KI-Systeme, die Entscheidungen unterstützen, sollten nicht nur eine Empfehlung ausgeben, sondern die Empfehlung als eine mögliche Einschätzung mit expliziter Unsicherheit darstellen. Das reduziert den Anchoring-Effekt. Und die Übergabe an die entscheidende Person sollte immer mit einer expliziten Zuweisung der Verantwortung verbunden sein.

Verschiebung 3: Die vier Rationalitätsfragen als KI-Governance-Instrument

Vor der Einführung eines KI-Systems in Entscheidungsprozesse lohnt es sich, vier Fragen strukturiert zu beantworten:

- Woher wissen wir das, was das KI-System empfiehlt? Auf welchen Daten und Modellen beruht die Empfehlung? Wer hat das validiert?
- Welches Interesse hat die Wissensquelle? Wer hat das System entwickelt und mit welchem Optimierungsziel? Entspricht dieses Ziel unserem?

- Unter welchen Voraussetzungen ist das behauptete Wissen gültig? Für welche Fälle wurde das System trainiert? Gilt das für unsere Fälle?
- Wie lauten die Definitionen der zentralen Begriffe? Was versteht das System unter „Geeignet“, „Risiko“, „Empfehlung“? Deckt sich das mit unserem Verständnis?

Diese Fragen sind nicht technisch. Sie sind Governance-Fragen. Und sie können nur von Menschen beantwortet werden, die das System, das Domänenwissen und die Unternehmensverantwortung zusammenbringen.

DIE DREI KERNFRAGEN FÜR JEDES KI-GESTÜTZTE ENTSCHEIDUNGSSYSTEM

- Wer kann diese KI-Empfehlung mit eigenem Domänenwissen prüfen — und ist das tatsächlich die Person, die formal die Aufsicht trägt?
- Wer weiß im Fehlerfall, was zu tun ist — nicht formell, sondern wirklich?
- Ist im Systemdesign sichtbar gemacht, dass die KI eine Empfehlung gibt und kein Urteil — und dass die Entscheidung beim Menschen liegt?

Eine Schlussbemerkung

Der EU AI Act ist ein ernsthafter Versuch, Verantwortlichkeit in KI-Systeme einzuschreiben. Aber Regelwerke können nicht die kognitiven Mechanismen regeln, die Menschen dazu bringen, Verantwortung abzugeben. Das können nur Organisationen selbst — durch die Art, wie sie KI-Systeme gestalten, einführen und in Entscheidungsprozesse einbetten.

Zukunft ist eine Entscheidung. Das gilt auch für die Frage, ob KI menschliche Entscheidungen ergänzt oder ersetzt. Diese Entscheidung trifft nicht das KI-System. Sie trifft jede Organisation, die KI einsetzt — durch das Design ihrer Prozesse, die Klarheit ihrer Verantwortlichkeiten und die Ehrlichkeit ihrer Governance.

Ein KI-System, das empfiehlt, verändert die Entscheidungsstruktur der Organisation. Wer diesen Eingriff nicht explizit gestaltet, lässt ihn implizit geschehen. Beide Varianten sind Entscheidungen. Nur eine davon ist eine bewusste.

Referenzen

- [1] Analyses of Social Issues and Public Policy (2025): AI as moral cover: How algorithmic bias exploits psychological mechanisms to perpetuate social inequality. Wiley. doi.org/10.1111/asap.70031
- [2] Carnat, I. (2024): Human, All Too Human: Accounting for Automation Bias in Generative Large Language Models. International Data Privacy Law. doi.org/10.1093/idpl/ipae018
- [3] Coeckelbergh, M. (2020): The Problem of Many Hands and the AI Responsibility Gap. In: AI Ethics. Springer.
- [4] Europäische Kommission (2024): Verordnung (EU) 2024/1689 – EU AI Act, insbesondere Art. 14 (Human Oversight). EUR-Lex.
- [5] Frontiers in Psychology (2023): Check the box! How to deal with automation bias in AI-based personnel selection. doi.org/10.3389/fpsyg.2023.1118723
- [6] Horowitz, M.C. & Kahn, L. (2024): Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts. International Studies Quarterly 68(2), sqae020. doi.org/10.1093/isq/sqae020
- [7] Journal of Business Ethics (2025): The Bright Side of AI in Marketing Decisions: Collaboration with Algorithms Prevents Managers from Violating Ethical Norms. Springer. doi.org/10.1007/s10551-025-06083-w
- [8] Journal of Computer-Mediated Communication (Oxford, 2022): How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability. doi.org/10.1093/jcmc/zmac029
- [9] Journal of Public Administration Research and Theory (Oxford, 2023): Human–AI Interactions in Public Sector Decision Making: Automation Bias and Selective Adherence. doi.org/10.1093/jpart/muac033
- [10] Kücking, F. et al. (2024): Automation Bias in AI-Decision Support: Results from an Empirical Study. Studies in Health Technology and Informatics 317, 298–304. PubMed 39234734. doi.org/10.3233/SHTI240871
- [11] Preprints.org (Nov. 2025): The Impact of AI-Based Decision Support Systems on Human Agency in Governance. preprints.org/manuscript/202511.1534
- [12] Scientific Reports (2023): Humans inherit artificial intelligence biases. Nature. doi.org/10.1038/s41598-023-42384-8
- [13] Springer AI & Society (2025): Exploring automation bias in human–AI collaboration: a review and implications for explainable AI. PRISMA 2020, 35 Studien. doi.org/10.1007/s00146-025-02422-7
- [14] Skitka, L.J. et al. (2000): Accountability and automation bias. International Journal of Human-Computer Studies 52(4), 701–717. (Grundlagenstudie zur Verantwortungsdiffusion, zitiert in: Frontiers 2023)