

KI-RESILIENZ IM JOURNALISMUS

Framework und Handlungsempfehlungen für mehr
Vertrauen, Verantwortung und digitale Souveränität



Whitepaper

Dr. Leif Kramp

Dr. Stephan Weichert

KI-Resilienz im Journalismus

Framework und Handlungsempfehlungen für mehr
Vertrauen, Verantwortung und digitale Souveränität

Vorbemerkung

Künstliche Intelligenz verändert den Journalismus nicht schrittweise, sondern strukturell. Sie greift in Recherche, Produktion und Distribution ein – und verschiebt damit Macht, Sichtbarkeit und Haftung. Wer heute pauschal über Tools redet, redet am eigentlichen Problem vorbei. Es geht *nicht* um Effizienzgewinne oder einzelne Anwendungsfälle. Es geht darum, ob und wie Journalismus unter KI-Bedingungen bestehen kann.

Denn mit KI steht mehr auf dem Spiel als Arbeitsabläufe: Vertrauen erodiert, wenn Inhalte ohne klar erkennbare Absender zirkulieren; Zuständigkeiten diffundieren, wenn Entscheidungen algorithmisch vorbereitet werden; Beziehungen zum Publikum werden porös, wenn neue KI-Interfaces Medienmarken verdrängen. Die zentrale Herausforderung lautet: *Wie bleibt Journalismus entscheidungssicher, zurechnungsfähig und glaubhaft in einer Kommunikations- und Medienumgebung, die zunehmend von KI geprägt ist?*

Dieses Whitepaper antwortet darauf nicht mit Tool-Listen oder Technikfolklore. Es entwickelt ein praxisnahes *KI-Resilienz-Framework für Redaktionen und Medienschaffende*. Zentral sind dabei Verantwortungszuweisungen, verbindliche *Human-in-the-Loop*-Prozesse, Transparenz gegenüber dem Publikum, Schutz von Daten und Quellen, überprüfbare Verifikation, ethische Leitplanken und systematischer Kompetenzaufbau durch Weiterbildung. KI ist dabei kein Selbstzweck, sondern ein Instrument, das mit Bedacht, ethisch sauber und organisational klug eingebunden werden muss.

KI-Resilienz verstehen wir als professionelle Kernkompetenz. Sie bedeutet nicht Verweigerung oder Widerstand gegen digitale Technologie, sondern beschreibt die Qualität von Redaktionen, unter den Bedingungen algorithmischer Zwänge souverän zu handeln – statt sich von technologischer Innovation treiben zu lassen oder in nervöse Abwehrreflexe zu verfallen. Dieses Whitepaper richtet sich an Redaktionen und freie Medienleute, die KI nutzen wollen, ohne Vertrauen zu verspielen. Es richtet sich an die Medienpolitik, die neue regulatorische Rahmen setzen muss, damit öffentliche Kommunikation nicht durch Plattformlogiken zersetzt wird. Und es richtet sich an die Zivilgesellschaft, die Resilienz als demokratischen Aushandlungsprozess begreifen muss.

KI-Resilienz ist weder Kampf- noch Modebegriff. Sie ist gelebte Praxis und setzt klare Zuständigkeiten, definierte Prüfprozesse, dokumentierte Eingriffe und transparente Kommunikation gegenüber Dritten voraus. Ob KI zum natürlichen Feind des Journalismus wird, entscheidet sich letztlich an ihrer Einbettung in redaktionelle Strukturen und Standards. Dafür legen wir mit diesem Whitepaper einen konkreten Gestaltungsvorschlag vor.

Die Autoren – Dr. Leif Kramp & Dr. Stephan Weichert, Februar 2026

Inhaltsverzeichnis

Vorbemerkung.....	2
Abbildungsverzeichnis.....	5
1. Einleitung: Warum „KI-Resilienz“ jetzt?	6
1.1 Journalismus am Kipppunkt.....	6
1.2 Journalistische Begegnungs- und Live-Formate als Gegentrend.....	7
1.3 Die Schizophrenie der Plattformabhängigkeit	9
1.4 Vertrauensbruch im digitalen Raum	12
1.5 Leitthese und Einordnung des Whitepapers.....	13
2. KI im Journalismus: Historische Linien, Projektionen, Brüche	15
2.1 Von Mythen, Maschinen und Missverständnissen: Der „ELIZA-Effekt“	15
2.2 Die „gezackte Grenze“ – warum KI leistungsfähig und dumm zugleich ist	18
2.3 Machtfragen: Plattformen, Modelle und neue Gatekeeper	22
2.4 Journalismus zwischen Effizienzgewinn und Identitätsverlust.....	24
3. Resilienz als Schlüsselkonzept im Journalismus	27
3.1 Vom KI-Bruch zur Digitalen Resilienz als Orientierungsrahmen.....	27
3.2 Resilienz begrifflich klären: Bewahren, Lernen, Anpassen	29
3.3 Drei Ebenen journalistischer Resilienz	31
3.4 Resilienz im Alltag: Spannungsfelder, Routinen und organisationale Praxis.....	33
3.5 Resilienz: professionelle Selbstvergewisserung im KI-Zeitalter	45
4. Leitlinien, Regelwerke und Positionspapiere	49
4.1 Warum Leitlinien? Normative Antworten auf strukturelle Unsicherheit.....	49
4.2 Internationale KI-Leitlinien und Orientierungsrahmen im Journalismus	50
4.3 Deutsche Leitlinien und Positionspapiere.....	53
4.4 Zwischen normativer Orientierung und erlebter Praxis	57
4.5 Leitlinien jenseits des Journalismus	63
4.6 Von normativer Selbstbindung zu organisationaler Handlungsfähigkeit	65

5. Medienpolitik und Regulierung: Zwischen Förderung und Kontrollillusion.....	66
5.1 Von redaktioneller Selbstbindung zu öffentlicher Rahmensetzung	66
5.2 Der EU AI Act und der Journalismus	68
5.3 Public-Interest-KI als gemeinwohlorientierte Ermöglichungsstrategie	69
5.4 Deutschland als Anwendungsfall.....	77
5.5 Konfliktlinien und Zielkonflikte: Freiheit, Schutz und Innovationsfähigkeit	79
5.6 Regulierung setzt Grenzen – Resilienz entsteht in der Praxis	82
6. KI-Framework für Redaktionen – Kein Toolkit für die Medienpraxis!.....	84
6.1 Warum Frameworks? Orientierungswissen statt Tool-Partys.....	84
6.2 Das KI-Resilienz-Framework für Redaktionen	89
6.3 Das Framework im Praxistest: Best Practices und Warnzeichen.....	106
6.4 Von der lernenden Redaktion zur gesellschaftlichen Resilienz.....	118
7. Ein neuer Gesellschaftsvertrag: KI-Resilienz als öffentliche Aufgabe	121
7.1 Warum ein neuer Gesellschaftsvertrag notwendig ist	121
7.2 Der normative Kern: Resilienz als Lernfähigkeit, nicht als Robustheit	123
7.3 Pflichten des Journalismus: Neue professionelle Zumutungen	125
7.4 Konsequenzen und Ableitungen nach Verbundebenen.....	126
7.5 Die Rolle des Staates: Ermöglicherin statt Technokratin	129
7.6 Macht, Asymmetrie und öffentliche Orientierung.....	130
7.7 Schlussgedanke: KI-Resilienz als demokratische Infrastruktur	131
Endnoten	132

Abbildungsverzeichnis

Abbildung 1: Tool- vs. Strukturperspektive auf den Umgang mit KI im Journalismus.....	7
Abbildung 2: Der „Weimatar“ des BKM Wolfram Weimer.....	10
Abbildung 3: Die virtuelle albanische Ministerin „Diella“.....	11
Abbildung 4: KI im Wandel der Zeit.....	17
Abbildung 5: Die gezackte Grenze der KI-Leistungsfähigkeit	20
Abbildung 6: Herausforderungen Digitaler Resilienz als Mehrebenen-Modell	32
Abbildung 7: Typologie journalistischer KI-Resilienz	34
Abbildung 8: Journalismus als vermittelnde Praxis digitaler Öffentlichkeit	47
Abbildung 9: Weiterentwicklung der KI-Richtlinien des BR von 2020 bis 2024.....	55
Abbildung 10: Sorgen in Bezug auf die fortschreitende Verbreitung von KI	59
Abbildung 11: Künftig stärker geforderte Rollen im Journalismus.....	60
Abbildung 12: Allgemeine Stimmung im eigenen Arbeitsumfeld	61
Abbildung 13: Von der Aufmerksamkeitsökonomie zur Vertrauensökonomie	67
Abbildung 14: KI-Resilienz im Journalismus als geteilte Gestaltungsaufgabe	81
Abbildung 15: Digitale Souveränität im Zusammenspiel.....	86
Abbildung 16: Das KI-Resilienz-Framework für Redaktionen	94
Abbildung 17: Redaktionelle KI-Governance als lernender Steuerungskreislauf	119

1. Einleitung: Warum „KI-Resilienz“ jetzt?

Worauf dieses Kapitel Antworten gibt:

1. Warum digitale Technologie- und Plattformabhängigkeit Redaktionen zugleich neue Chancen eröffnet und neue Verwundbarkeiten schafft.
2. Wie KI als Katalysator Vertrauenskrisen verschärfen, Polarisierung befeuern und bestehende Bruchlinien in der Öffentlichkeit sichtbar machen kann.
3. Wie wir mit einem praxisnahen Framework Redaktionen, Medienpolitik und Zivilgesellschaft bei Entscheidungen und Prioritäten im KI-Wandel unterstützen wollen.
4. Warum dieses Whitepaper kein Technik- oder Ethikpapier ist, sondern als Orientierungskompass für Resilienz, Beziehungsgestaltung und digitale Souveränität dient.

1.1 Journalismus am Kipppunkt

Der Journalismus steht an einem Kipppunkt, der sich eher still und heimlich angekündigt hat – und jetzt mit voller Wucht in das Bewusstsein vieler Redaktionen dringt. Künstliche Intelligenz ist kein Experimentierfeld mehr, kein futuristisches Szenario, kein bloßer Showcase. KI ist zur realen Grundinfrastruktur der digitalen Öffentlichkeit geworden – oder entwickelt sich zumindest immer mehr dorthin: Sie wirkt im Hintergrund, ordnet Informationen, verschiebt Bedeutungen, flirtet mit der Realität, setzt politische Agenden.

Lange haben wir geglaubt, KI sei ein reines, funktionales Werkzeug wie jedes andere Tool: nützlich, effizient, vielleicht sogar revolutionär, aber eben beherrsch- und steuerbar. Wie ein digitaler Taschenrechner – oder ein Algorithmus, der einfach ein bisschen schlauer wirkt als die Version davor. Doch während wir aktuell in Redaktionen noch darüber streiten, ob KI unsere Arbeit erleichtern oder unsere Produkte eleganter machen kann, hat sich der eigentliche Umbruch längst vollzogen: KI wird im Hintergrund zur heimlichen Arbeits- und Entlastungsinfrastruktur, die viele Medienschaffende bereits täglich nutzen, ohne dies an die große Glocke zu hängen. KI ist außerdem schleichend zu einer gigantischen Projektionsfläche geworden – für Ängste und Hoffnungen Einzelner, aber auch für Effizienzfantasien vieler Verlage.

Wir erleben nichts weniger als eine *digitale Zäsur*, in der es längst nicht mehr um technische Evolution geht als vielmehr um die Gretchenfrage, ob sich journalistische Inhalte in den kommenden Jahren noch behaupten können, ob journalistische Arbeit überhaupt noch sichtbar bleibt und ob Menschen sie weiterhin zu schätzen wissen – und bereit sind,

dafür zu bezahlen –, während vornehmlich US-amerikanische Plattformkonzerne und KI-Sprachmodelle (LLMs) immer tiefer und direkter in die Herstellung von Öffentlichkeit eingreifen.

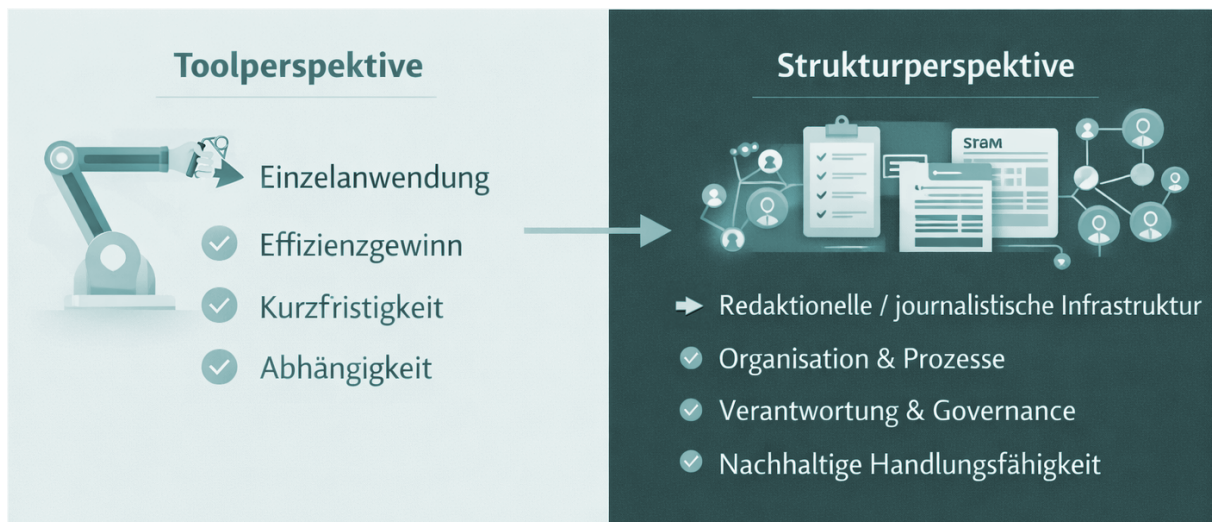


Abbildung 1: Tool- vs. Strukturperspektive auf den Umgang mit KI im Journalismus (eigene Darstellung)

Seit Jahren leben Redaktionen in einer *strukturell riskanten Doppelabhängigkeit*: von Plattformalgorithmen, die Sichtbarkeit und Reichweite steuern – und von Technologien, die immer stärker direkt in die Produktionslogik des Journalismus eingreifen. Beides bedroht die Integrität und Souveränität der journalistischen Praxis. Und KI – das ist der Fokus des vorliegenden Whitepapers – fügt jeder dieser Strömungen eine neue Dynamik hinzu: Sie liefert Effizienzgewinne und Entlastungsversprechen qua Automatisierung, vertieft aber zugleich hergebrachte Gräben: Sie droht journalistische Inhalte zu entwerten, das Publikum zu fragmentieren, und auch gemeinsame Realitäten zu erodieren. Die kapitalistische Digitalökonomie mit ihrem Ursprung im Silicon Valley hat diese Entwicklungen lange vorbereitet. KI macht sie nun unübersehbar – und verschiebt den Journalismus zunehmend in eine *Vertrauensökonomie*.

1.2 Journalistische Begegnungs- und Live-Formate als Gegentrend

Es ist daher kein Zufall, dass es innerhalb wie außerhalb von Redaktionen wieder sehr viel stärker um die *Menschlichkeit journalistischer Arbeit* geht, das persönliche Gespräch, die Präsenz vor Ort, um Begegnung – gemeinhin mit dem Begriff „Dialog- oder Community-Journalismus“ tituiert.¹ Darunter werden nicht nur journalistische Live-Auftritte

verstanden, sondern auch Erzählcafés, Theaterstücke. Bühnenshows, Poetry bzw. Reporter Slams oder andere Begegnungs- und Live-Formate verstanden, mit denen Journalist:innen bewusst den Kontakt zur lokalen Öffentlichkeit suchen und das direkte Gespräch aufnehmen. Die konsequenteste Ausprägung dieses Ansatzes sind Begegnungsorte oder Akademien, die sich als „Orte des Journalismus“² verstehen – als manifeste, zugängliche Anlaufstellen für journalistische Debatten, Austausch und Öffentlichkeit. In jüngerer Zeit entstehen in diesem Zusammenhang auch Projekte wie „Media Rewilding“ von Alexander von Streit, das gezielt neue Formen der lokalen Begegnung zwischen Journalismus und Publikum erprobt (<https://media-rewilding.de>).

Insofern lässt sich ein solcher Dialog- oder Live-Journalismus auch als bewusster Gegen-trend zur fortschreitenden Synthetisierung digitaler Öffentlichkeit verstehen. Denn KI entfaltet eine *doppelte Dynamik*: Sie treibt Prozesse einerseits weiter an, die ohnehin längst im Gange sind, andererseits legt sie die strukturelle Verwundbarkeit moderner Medien offen: *Deepfakes*, synthetische Stimmen, automatisierte Propaganda – all das ist kein bloßes technisches Rauschen, sondern greift das Kernversprechen journalistischer Arbeit – und damit das in die mediale Wahrnehmung der Welt an: *Glaubwürdigkeit*. In einer Zeit, in der sich Wahrheit und Fälschung nur noch nuanciert unterscheiden, geraten selbst etablierte Medien ins Wanken. Beim Publikum kann sich – verstärkt durch den zunehmenden KI-Einsatz im Journalismus – schnell das Gefühl einstellen, dass ihm bei der Mediennutzung die Realität entgleitet, Nutzer:innen erleben also einen ernsten „Authentizitätskollaps“³ bzw. *Reality Slip*:⁴ Was kann man noch glauben? Was ist noch verlässlich?

Hinzu kommen teils widersprüchliche Narrative, die den öffentlichen Diskurs über KI polarisieren. Auf der einen Seite stehen die Heilsversprechen der Tech-Eliten, die KI als quasi-erlösende Zukunftstechnologie inszenieren. Auf der anderen Seite eine Apokalypse-Rhetorik, die das Ende menschlicher Autonomie beschwört. Dazwischen verläuft eine stille, aber entscheidende Grauzone journalistischer Selbstbehauptung: jener Bereich, in dem Resilienz entstehen oder aber verloren gehen kann. Es nimmt kaum Wunder, dass selbst KI-Pioniere wie der britische Informatiker und Nobelpreisträger Geoffrey Hinton inzwischen vor den menschenverachtenden Konsequenzen einer entfesselten KI warnen.⁵ Ebenso bezeichnend ist, wie schnell die Euphorie – auch in den Medien – in Alarmismus umschlagen kann, sobald deutlich wird, dass die neuen Systeme nicht nur Werkzeuge sind, sondern zusehends zu *Machtinstrumenten in einem globalen Kulturkampf* werden.⁶

Dieses *Ambivalenz-Narrativ* – Erlösung und Untergang im Doppelpack – hat eine klare Folge: Es überlagert den nüchternen Blick, den wir eigentlich auf digitale Technologie richten sollten, um uns nicht in unserer Urteilsfähigkeit einschränken zu lassen. KI

erscheint dann als metaphysische Kraft und weniger als das, was sie tatsächlich ist – ein statistisches System, das auf menschlicher Arbeit, extrahierten Daten und aktuell schier unermesslichen Finanzressourcen beruht, aber auch zielgerichtet eingesetzt werden kann, um weltwirtschaftliche Vormachtstellungen zu zementieren. Genau an diesem Punkt setzt der Begriff der *KI-Resilienz* an. Wie wir zeigen werden, richtet er sich gegen die Überhöhung der Maschine als hyperkomplexes, oft bewusst opakes System, aber ebenso gegen die lähmende Angst, die diese Technologie auslösen kann. KI-Resilienz beschreibt die Fähigkeit, sich im Spannungsfeld von Beschleunigung und Entgrenzung zu behaupten: kritisch und reflektiert, zugleich aber anpassungs- und handlungsfähig.

Für Redaktionen bedeutet das, sich nahezu täglich mit neu entstehenden Widersprüchen auseinanderzusetzen. KI kann Texte zusammenfassen, Bilder erzeugen, Fakten prüfen, Daten durchsuchen – und doch bleibt die zentrale Frage offen: Wessen Urteil zählt? Wer hat das letzte Wort? Und wann gilt ein Inhalt überhaupt als gesichert?

Wenn Verantwortung zwischen Menschen und Maschinen verschwimmt, wird die journalistische Selbstwirksamkeit durchlässiger. Was früher ein klarer Auftrag war – prüfen, ordnen, erklären –, wird heute von Systemen überlagert, die zwar schnell, aber nicht verantwortlich ‚handeln‘, und keineswegs fehlerfrei, denn:

- *Wer ist zuständig, wenn ein KI-generierter Beitrag Fehler oder Desinformationen enthält?*
- *Wer entscheidet, was als vertrauenswürdig gilt?*
- *Und was bedeutet es für eine Redaktion, wenn das Publikum zunehmend auf Chatbots zurückgreift, deren Ergebnisse nicht – oder nicht zwingend und jedenfalls allzu selten – auf journalistisch geprüften Informationen basieren?*

1.3 Die Schizophrenie der Plattformabhängigkeit

Dabei wächst die Abhängigkeit von Journalismus und Öffentlichkeit von Plattformen nicht nur immer weiter, sondern sie nimmt eine völlig neue Qualität an: Google hat inzwischen *Zero-Click-Suchen* eingeführt, bei denen Antworten direkt aus KI-Modellen stammen, ohne dass Nutzer überhaupt noch auf deren (journalistische) Quellen gelangen.⁷ Die ökonomische Grundlage des Journalismus droht sich damit weiter zu verschieben – weg vom publizistischen Inhalt, hin zur technischen KI-Infrastruktur. Und während Tech-Konzerne längst den Bau gigantischer KI-Serverfarmen wie des „Stargate“-Projekts in Texas bauen, ringen Redaktionen mit Betriebsvereinbarungen über Toolfreigaben oder mit der Frage, ob

Journalist:innen künftig noch das ‚Prompting‘ lernen müssen. Die Kluft zwischen digitalem Zukunftspathos und redaktioneller Realität könnte kaum größer sein.



Abbildung 2: Der „Weimatar“ des Beauftragten der Bundesregierung für Kultur und Medien Wolfram Weimer⁸

Die politischen Implikationen reichen hingegen längst über Verlagsgebäude hinaus und sind nicht weniger dramatisch: Die laufende Klage der „New York Times“ gegen OpenAI/ Microsoft und Perplexity wegen mutmaßlicher Urheberrechtsverletzungen (der inzwischen viele anderen Medienhäuser gefolgt sind)⁹ und Debatten um KI-Fakes im US-Wahlkampf sowie ihr gehöriger Einfluss auf Wahlen weltweit zeigen,¹⁰ wie stark demokratische Öffentlichkeit inzwischen unter Druck geraten ist. KI ist längst ein *politisches Machtinstrument* – und wer sie einsetzt, kontrolliert oder reguliert, entscheidet zunehmend darüber, wie Debatten geführt werden und wer Wahlen gewinnt. Dass nicht nur Albanien Ministerpräsident erstmalig eine KI namens „Diella“ (Abb. 3) zur virtuellen Ministerin ernennt,¹¹ sondern fast zeitgleich der deutsche Medien- und Kulturstatsminister Wolfram Weimer im September 2025 seine KI-Version „Weimatar“ öffentlich präsentiert,¹² während zugleich Expertinnen und Experten zur Vorsicht im Umgang mit KI-Doppelgängern warnen, illustriert die Ambivalenz der KI-Bewegung: Wir bewegen uns tastend zwischen technischer Spielerei und bevorstehender Vertrauenskrise, die vor allem das politische System belastet – wie sich in den USA schon seit einigen Jahren beobachten lässt.¹³



Abbildung 3: Die virtuelle albanische Ministerin „Diella“ in einem Videopost der französischen Nachrichtenagentur AFP auf X¹⁴

In dieser Gemengelage reicht es nicht mehr aus, lediglich KI-Tools einzuführen oder neue Leitlinien zu veröffentlichen. Was Journalist:innen brauchen, ist ein *klarer, gemeinsamer Orientierungsimpuls*: eine Haltung, die jenseits von Allmachtsfantasien und Apokalypse-Erzählungen bleibt; ein resilientes Framework, das – wie wir später zeigen – Vertrauen, Verantwortung und bewusste Beziehungsarbeit in den Mittelpunkt des Umgangs mit KI stellt. Denn Resilienz ist keine defensive Handlung, sondern die Fähigkeit, Beziehungen aktiv und reflektiert zu gestalten – zwischen Menschen, Medien und Maschinen.

Genau an diesem Punkt setzt dieses Whitepaper an. Es formuliert KI-Resilienz als journalistische Schlüsselkompetenz und als demokratische Praxis. Es bietet einen universellen Resilienzrahmen für Redaktionen, zeigt konkrete Handlungsspielräume für Medienpolitik auf und eröffnet Perspektiven für eine Zivilgesellschaft, die digitale Souveränität nicht nur fordern, sondern auch praktisch zurückgewinnen kann. KI ist in diesem Verständnis kein Technik-Hype, der über uns hereinbricht, sondern ein sozialpolitisches Projekt – mit Risiken und Potenzialen, aber immer auch mit Verantwortlichkeiten und der Notwendigkeit, unsere zivilgesellschaftlichen Gestaltungsrolle zu behaupten. Der Verlauf dieses massiven Wandels hängt also davon ab, *ob und wie wir ihn bewusst steuern*. Das Stadium der reinen Tool-Debatte – *Was können KI-Tools leisten? Kann und soll KI journalistische Inhalte produzieren?* – liegt aus unserer Sicht längst hinter uns.

Die eigentliche Frage lautet daher nicht, was KI kann, sondern was sie mit dem sozialen Gefüge der Öffentlichkeit macht – und wie Redaktionen in diesem neuen Ökosystem handlungsfähig bleiben können.

1.4 Vertrauensbruch im digitalen Raum

Schon lange vor der aktuellen KI-Welle waren Redaktionen abhängig von Plattformen, deren Regeln und Geschäftsmodelle sie kaum beeinflussen konnten. Google entscheidet seit mehr als zwei Jahrzehnten über die Sichtbarkeit, Facebook über Reichweiten, TikTok über kulturelle Relevanz. Mit generativer KI beginnt jedoch eine neue Phase dieser Abhängigkeit. Nicht mehr die Nutzenden entscheiden primär, welche Inhalte sie anklicken, sondern KI-Systeme beantworten deren Fragen mittels eines Chat-Interfaces. *Google Zero* – siehe oben – ist dafür das prägnanteste Beispiel: Antworten entstehen direkt im KI-Chat-Bot, ohne dass Originalquellen transparent sichtbar werden.

Damit werden KI-Systeme zu *Gatekeepern zweiter Ordnung*. Sie schaffen nicht nur Zugänge zu Informationen, sie produzieren Informationen selbst mittels eines (in der Regel intransparenten) mit unzähligen Daten gefütterten Sprachmodells. Die Ambivalenz ist offensichtlich: KI verspricht Entlastung, Automatisierung und Effizienz. Gleichzeitig entsteht eine infrastrukturelle Macht, die sich journalistischer Kontrolle weitgehend entzieht. Redaktionen laufen Gefahr, zu bloßen Daten- und Trainingslieferanten für KI-Modelle zu werden, die später mit ihren eigenen Inhalten konkurrieren. Die Wertschöpfung verschiebt sich damit endgültig weg vom journalistischen Inhalt – hin zu den Systemen, die diese Inhalte aggregieren, imitieren und industriell verwerten.

Parallel dazu zerfällt das, was die demokratische Öffentlichkeit über Jahrzehnte zusammengehalten hat: Vertrauen. Die großen Krisen der vergangenen Jahre – Pandemie, Kriege, populistische Bewegungen – haben das Medienvertrauen weltweit unter Druck gesetzt. KI kann in diesem ohnehin hoch aufgeladenen Klima wie ein Teilchenbeschleuniger wirken. Je perfekter synthetische Bilder und Stimmen werden, desto herausfordernder wird die Unterscheidung zwischen Fakt und Fiktion.

Der dadurch erzeugte *Reality Slip* – siehe oben – ist also ganz real: Menschen erleben eine mediale Welt, die sich immer weniger mit ihren eigenen Alltagserfahrungen deckt. Deepfakes, synthetische Statements, KI-optimierte Propaganda – all das untergräbt nicht nur Informationssicherheit, sondern die grundlegende Gewissheit, dass Wirklichkeit überhaupt noch verlässlich auffindbar ist. Für den Journalismus bedeutet das einen *Verlust an Beziehungskapital*. Denn Vertrauen entsteht nicht durch perfekte Inhalte, sondern durch nachvollziehbare, wiedererkennbare Beziehungen. Wenn diese Beziehungen

zunehmend durch künstlich erzeugte Konkurrenzprodukte ersetzt werden, verliert der Journalismus sein wichtigstes Gut: Glaubwürdigkeit als gesellschaftliche Legitimation.

KI verstärkt nun diese Entwicklung drastisch: *Je perfekter die Imitation, desto schwieriger die professionelle Distinktion.* Und je schneller die Systeme antworten, desto weniger Raum bleibt für das Entschleunigungsmoment journalistischer Standards wie Prüfung, Kontextualisierung oder Relevanzbewertung. Journalist:innen verlieren als Torwächter für den geordneten Zugang zu Informationen an Bedeutung – nicht, weil ihre Arbeit schlechter würde, sondern weil die infrastrukturelle Umgebung schriller, schneller und glatter operiert. Genau hier lauert die zentrale Gefahr: Journalismus könnte schon bald nicht mehr als notwendige (systemrelevante) Infrastruktur wahrgenommen werden, sondern nur noch als eine Option unter vielen.

Für Redaktionen ist das kein technisches Detail, sondern ein *resilienzstrategisches Kernproblem*. Es geht nicht darum, ob KI intelligenter Inhalte produzieren kann, sondern darum, ob journalistische Unabhängigkeit in einer Architektur erhalten bleibt, die nicht von Journalist:innen, sondern von Coding-Teams, Plattformkonzernen und Investorengremien gestaltet wird. Die Frage lautet daher nicht, wie effizient Redaktionen die KI nutzen – sondern *inwiefern sie sich in einem System behaupten, dessen Performance sie nicht kontrollieren können.*

1.5 Leitthese und Einordnung des Whitepapers

Unsere folgenden Ausführungen verdeutlichen: Die meisten KI-Leitlinien, die heute existieren – ob von internationalen Medienhäusern, journalistischen Netzwerken oder Branchenverbänden – konzentrieren sich auf Transparenz, Kennzeichnung, Sicherheit und Datenethik (vgl. Kap. 4). All diese Aspekte sind wichtig, aber sie bleiben oft an der Oberfläche. Sie schaffen Regeln, aber keine Souveränität. Sie definieren *Do's und Don'ts*, aber keine Orientierung für jene Situationen, in denen die Tools technisch funktionieren – und der Mensch trotzdem überfordert ist (vgl. Kap. 2, 4).

Was unter anderem fehlt – und deshalb gibt es dieses Whitepaper – ist der ganzheitlich-ethische, aber auch der psychosozial-organisationale Blick, was landläufig unter dem Schlagwort der *Responsible AI*¹⁵ verhandelt wird (vgl. Kap. 2, 3, 6):

- *Wie verändert KI die Atmosphäre in Redaktionen?*
- *Wie verändert sie Teamdynamik, Verantwortungslinien, Entscheidungsprozesse?*
- *Wie verändert sie das Verhältnis zu Quellen, Fehlern, zum Publikum?*
- *Wie wirkt KI auf die Öffentlichkeit, die der Journalismus mitgestalten will?*

- *Und: Wie sehr stärken oder schwächen KI-Anwendungen die digitale Souveränität?*

Die zentrale Leerstelle ist aus unserer Sicht das, was wir *KI-Resilienz* nennen (vgl. Kap. 3). Sie wird bislang selten definiert, noch seltener operationalisiert (vgl. Kap. 6). Dabei entscheidet genau ihre Umsetzung darüber, ob Redaktionen KI souverän nutzen – oder von ihr getrieben werden.

Es fehlt somit ein praktisches Orientierungsmodell, das Redaktionen nicht überfordert, sondern langfristig stärkt. Ein Rahmen, der Widerstandskraft, Lernfähigkeit und Anpassungsbereitschaft zusammendenkt, statt sie gegeneinander auszuspielen. Und der die zentralen Funktionen des Journalismus schützt, ohne Innovation reflexhaft zu blockieren. Genau diese Leerstelle will dieses Whitepaper schließen.

KI-Resilienz bedeutet für uns deshalb drei miteinander verschränkte Dimensionen:

Erstens: Vertrauen nicht vorauszusetzen, sondern *aktiv aufzubauen* – gegenüber dem Publikum, innerhalb der Redaktion und im bewussten Umgang mit der Technologie selbst.

Zweitens: Verantwortung nicht zu delegieren, sondern *klar zu verorten*. KI kann unterstützen, Muster erkennen und Vorschläge machen – aber sie kann keine Urteile fällen und keine Wirklichkeit bewerten. Verantwortung bleibt menschlich, muss es bleiben.

Drittens: Beziehungen aktiv zu gestalten – zwischen Menschen, Medien und Maschinen. Resilienz ist dabei keine Abwehrreflexe, sondern *eine Form digitaler Mündigkeit*. Sie erkennt, dass nicht die Technik selbst das Problem ist, sondern der gesellschaftliche und organisationale Rahmen, in dem wir sie einsetzen.

Dieses Whitepaper will in diese Richtung orientieren –nicht im Sinne eines klassischen Leitfadens. Es liefert ein Framework, das sich an der Realität redaktioneller Prozesse orientiert, nicht an theoretischen Wunschvorstellungen. Es richtet sich an Medienhäuser, die KI einsetzen wollen, ohne ihre Glaubwürdigkeit zu verspielen. Es bietet Handlungsempfehlungen, die auch Freischaffende befähigen sollen, KI souverän, verantwortungsvoll und transparent einzusetzen. Es richtet sich ebenso an eine Medienpolitik, die mehr tun muss als zu regulieren – sondern ermöglicht, befähigt und Ressourcen schafft. Und es eröffnet Perspektiven für eine Zivilgesellschaft, die digitale Souveränität als zentrale demokratische Kompetenz erlebt.

Um KI-Resilienz im Hier und Jetzt zu praktizieren, müssen wir verstehen, wie weit die Geschichte dieser Verwerfungen zurückreicht. KI-Resilienz beginnt nicht mit ChatGPT, sondern mit Joseph Weizenbaum und dem sogenannten „ELIZA-Effekt“.

2. KI im Journalismus: Historische Linien, Projektionen, Brüche

Worauf dieses Kapitel Antworten gibt:

1. Warum der „ELIZA-Effekt“ sichtbar macht, inwieweit Menschen den Maschinen Bedeutung und Urteilskraft zuschreiben.
2. Wie sich die Entwicklung von frühen KI-Mythen bis zu ChatGPT als Massenphänomen historisch nachzeichnen lässt.
3. Wo sich heute Brüche zwischen Euphorie, Skepsis und institutioneller Überforderung im Umgang mit KI zeigen.
4. Welche Machtfragen sich durch Plattformen, Modelle und neue Gatekeeper stellen – von Big Tech über Dark Tech bis zu imperialen Logiken.
5. Was diese Entwicklungen konkret für redaktionelle Praxis, Verantwortung und Arbeitskultur bedeuten.

2.1 Von Mythen, Maschinen und Missverständnissen: Der „ELIZA-Effekt“

Die Geschichte der Künstlichen Intelligenz beginnt nicht mit Rechenzentren, Start-ups oder milliardenschweren Plattformen. Auch nicht mit einem großen Knall im Medienlabor. Sie beginnt mit einem Irrtum. Genauer gesagt: mit einer menschlichen Projektion.

Mitte der 1960er-Jahre entwickelte der Informatiker Joseph Weizenbaum am MIT ein simples Textprogramm mit dem Namen ELIZA.¹⁶ Das System war technisch banal: Es erkannte Schlüsselwörter und antwortete mit vorgefertigten Satzmustern. ELIZA verstand nichts. Es rechnete nicht, es dachte nicht, es hatte kein Bewusstsein. Und doch geschah etwas, das Weizenbaum selbst überraschte – und später zutiefst beunruhigte: Nutzer:innen begannen, ELIZA menschliche Eigenschaften zuzuschreiben. Sie vertrauten dem Programm persönliche Gedanken an. Sie führten Gespräche, als gebe es ein verständnisvolles Gegenüber. Nicht weil die Maschine intelligent war, sondern weil der Mensch bereit war, ihr diese Intelligenz zuzuschreiben.¹⁷

Was als Experiment begann, entwickelte sich rasch zu einer *kulturellen Schlüsselerfahrung*. Weizenbaum erkannte, dass nicht die Leistungsfähigkeit der Maschine das eigentliche Problem war, sondern das, was die Menschen daraus machen. Aus diesem Eindruck

heraus entstand der Begriff *ELIZA-Effekt* – ein Terminus,¹⁸ der sich später in der Fachliteratur, Technikfolgenabschätzung und Medienwissenschaft als nachträgliche Benennung genau dieses von Weizenbaum beschriebenen Phänomens einbürgerte und allgemeiner unter dem Begriff der *Antromorphisierung des Computers* diskutiert wird.¹⁹ Er beschreibt die Tendenz, dass wir Maschinen durchaus Verständnis, Empathie und Intentionalität zuschreiben, obwohl sie diese Fähigkeiten gar nicht besitzen:

„Indeed, the tendency to invest private feelings in a computer puzzled and concerned Weizenbaum, who worried that people’s internal reality might be replaced by that of the machine. Weizenbaum was also concerned by the extent to which computers ‘induce powerful delusional thinking in quite normal people’ and strengthened notions of human beings as machines, by which rationality became associated with calculation. This became known as the ‘ELIZA effect,’ the propensity for humans to ascribe understanding and intelligence to computer systems.“
(Berry 2023)

Dieser Effekt ist kein historisches Kuriosum. Er ist die Ursuppe all dessen, was wir heute im Umgang mit generativer KI nach einem einfachen Prinzip erleben: Je überzeugender eine Maschine kommuniziert, desto eher neigen Menschen dazu, sie für kompetent, authentisch und urteilsfähig zu halten. . Der „ELIZA-Effekt“ hilft also dabei zu erklären, warum Systeme auch dann als vertrauenswürdig gelten, wenn eigentlich klar ist, dass ihre Funktionsweise – vereinfacht gesagt – auf mathematischen Berechnungen beruht.

Weizenbaum zog daraus eine Konsequenz, die ihn zum Außenseiter seiner Zunft machte. Während viele seiner Kollegen KI als Fortschrittsmotor feierten, warnte er eindringlich davor, menschliche Entscheidungsprozesse an Maschinen zu delegieren: Nicht alles, was berechenbar sei, dürfe – so Weizenbaum – automatisiert werden, weil sich manche Herausforderungen, die Verantwortung, Urteil und moralische Abwägung erfordern schlicht der maschinellen Verarbeitung entzögen.²⁰

Heute wirkt diese keineswegs unbillige Warnung erstaunlich aktuell. Denn der „ELIZA-Effekt“ scheint längst zu skalieren: Aus einem einfachen Textprogramm ist eine globale Infrastruktur geworden. KI evoziert einen exponentiell gewachsenen Wust an digitalen Inhalten – sichtbar wie unsichtbar: Aus einer psychologischen Beobachtung wird ein struktureller Belastungstest für die demokratische Öffentlichkeit.

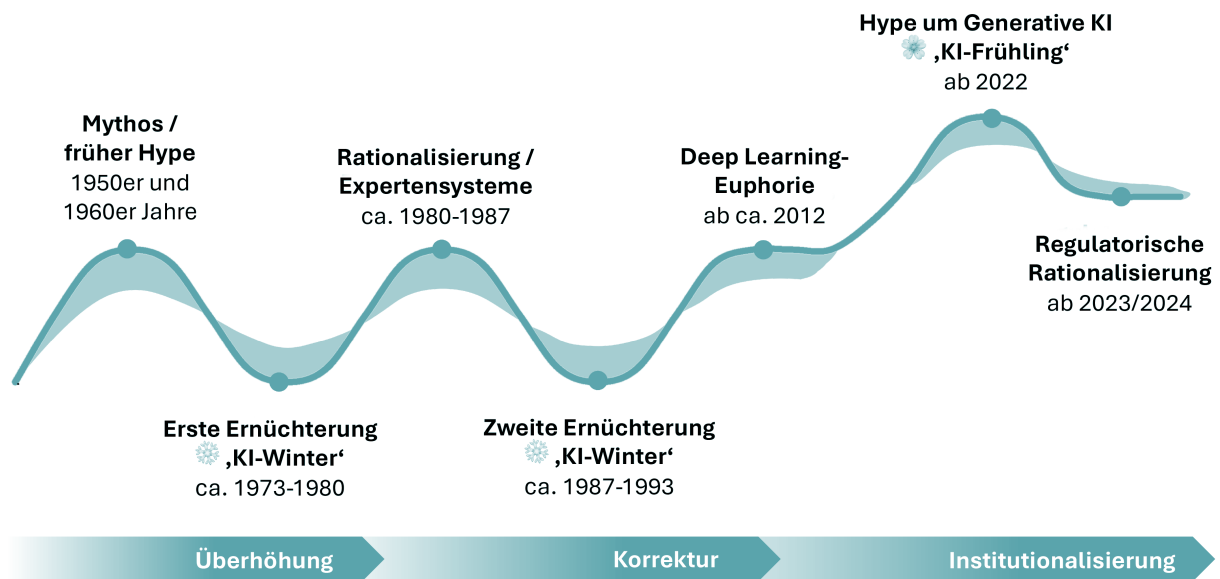


Abbildung 4: KI im Wandel der Zeit – Historische Zyklen von Mythos/Euphorie bis Neustart (eigene Darstellung)

Abbildung 4 fasst die zyklische Entwicklung von KI als wiederkehrende Abfolge von Erwartungsüberschuss, Ernüchterung und Neuorientierung zusammen. Der erste Ausschlag („Mythos“) steht für die frühe Phase symbolischer KI in den 1950er- und 1960er-Jahren – rund um ELIZA und science-fictionhafte Visionen denkender Maschinen. In dieser Phase wurde Maschinenintelligenz projektiv überhöht; die Erwartungen an schnelle Durchbrüche waren immens.

Darauf folgt die erste Ernüchterung, der sogenannte „erste KI-Winter“ (ca. 1973–1980). Der Lighthill-Report im Vereinigten Königreich und ähnliche Evaluierungen führen zu massiven Förderkürzungen. Viele Versprechen erweisen sich als nicht einlösbar, Investitionen gehen zurück, und die öffentliche Aufmerksamkeit sinkt deutlich.

In der anschließenden Phase der „Rationalisierung“ (ca. 1980–1987) verlagert sich der Fokus auf Expertensysteme: KI wird stärker regelbasiert, domänenspezifisch und industriell verwertbar gemacht. Die Technologie verliert ihren universalistischen Anspruch und wird als Spezialwerkzeug für klar definierte Anwendungen positioniert.

Diese Entwicklung mündet schließlich in einen zweiten „KI-Winter“ (ca. 1987–1993). Der Markt für Expertensysteme bricht teilweise ein, viele Systeme erweisen sich als teuer, unflexibel und wartungsintensiv. Die Erwartungen korrigieren sich abermals, weil die Technologie den hochgesteckten Versprechen nicht gerecht wird.

Mit dem Durchbruch datengetriebener Verfahren beginnt ab 2012 eine neue Boom-Phase: Die „Deep-Learning-Euphorie“ basiert auf großen Datenmengen, steigender Rechenleistung und erheblichen Fortschritten in Sprach- und Bilderkennung. KI wird nun nicht mehr als regelbasiertes System verstanden, sondern als lernfähige, statistische Infrastruktur.

Diese Dynamik verstärkt sich ab 2022 mit dem „Hype um Generative KI“, dem „KI-Frühling“: Große Sprachmodelle und multimodale Systeme erreichen – u.a. mit ChatGPT – breite gesellschaftliche Durchdringung: Die Technologie verlässt den Forschungs- und Unternehmenskontext und wird zum Massenphänomen. Die Erwartungen steigen erneut sprunghaft an – sowohl in wirtschaftlicher als auch in kultureller Hinsicht: KI wird zunehmend sozial angeeignet, Fortschritte in Rechenleistung, Datenverfügbarkeit und neuronalen Netzen erzeugen erneut Hype-Dynamiken – nun jedoch auf breiter gesellschaftlicher und wirtschaftlicher Basis.

Seit ca. 2023/24 zeichnet sich schließlich eine Phase „regulatorischer Rationalisierung“ ab. Mit Initiativen wie dem EU AI Act und einer intensivierten AI Governance-Debatte verschiebt sich der Diskurs von bloßer Leistungsfähigkeit hin zu Fragen von Verantwortung, Transparenz, Haftung und Machtasymmetrie. KI wird zunehmend institutionell eingebettet und normativ gerahmt.

In dieser Gesamtschau zeigt sich: Die Entwicklung der KI verläuft nicht linear, sondern in wiederkehrenden Wellen. Phasen technologischer Überhöhung werden regelmäßig durch Ernüchterung korrigiert und anschließend in stabilere, institutionell eingebettete Strukturen überführt. Der aktuelle Moment markiert daher weniger einen endgültigen Durchbruch als einen weiteren Übergang – von euphorischer Expansion hin zu gesellschaftlicher Einhegung und struktureller Integration.

Gerade vor diesem Hintergrund wird deutlich, warum der Umgang mit KI weniger eine Tool-Frage ist als eine *dauerhafte Resilienzaufgabe für den digitalen Journalismus*: Die eigentliche Herausforderung von KI liegt nicht in ihrer Rechenleistung, sondern in ihren Folgen für Vertrauen, Autorität und Beziehung. Nicht die Maschine entscheidet, wie mächtig sie wird – sondern die Art, wie ihr Bedeutung zugeschrieben wird.

2.2 Die „gezackte Grenze“ – warum KI leistungsfähig und dumm zugleich ist

Was ELIZA im stillen Kämmerlein vorführte, setzt sich heute im großen Maßstab fort. Der entscheidende Unterschied liegt nicht in der Logik, sondern im Maßstab. Moderne KI-

Systeme simulieren kein Gespräch mehr im Labor, sondern strukturieren Kommunikation im Alltag. Sie beantworten Fragen, fassen Texte zusammen, sortieren Informationen, schreiben E-Mails und moderieren Debatten. Damit sind sie von der Simulation zur Infrastruktur geworden.

Dabei bleibt der Kern unverändert: Diese Systeme wissen nichts. Sie verstehen nicht, was sie sagen. Sie berechnen Wahrscheinlichkeiten auf Basis gigantischer Datenmengen. Sie suggerieren Anschlussfähigkeit, erzeugen aber keine Erkenntnis. Doch je flüssiger, präziser und selbstbewusster ihre Antworten klingen, desto stärker entsteht der Eindruck von Kompetenz. Genau hier setzt die zentrale Verwechslung an: Verfügbarkeit wird mit Verlässlichkeit verwechselt, Plausibilität mit Wahrheit.

Was früher als technisches Experiment galt, ist heute fest in alltägliche Routinen eingebettet: Suchmaschinen, soziale Netzwerke, Redaktionssysteme und Content-Management-Tools – überall arbeiten bereits KI-gestützte Prozesse im Hintergrund. Sie strukturieren Aufmerksamkeit, priorisieren Themen, glätten Sprache und beschleunigen Entscheidungen. Der Journalismus nutzt diese Systeme zunehmend nicht mehr nur als professionelle Prothese, sondern als integralen Bestandteil seiner Produktionslogik.

Damit verändert sich auch die Rolle des Journalismus. Er agiert nicht mehr allein als vermittelnde Instanz zwischen Ereignis und Publikum, sondern bewegt sich zunehmend innerhalb einer technischen Architektur, die eigenen Regeln und Kausalitäten folgt. Die Maschine wird dabei nicht zum Akteur im klassischen Sinn – aber sie wird zum *strukturellen Mitspieler im Prozess öffentlicher Meinungsbildung*.

Gerade darin liegt eine auch für Journalist:innen mulmige Ambivalenz: KI kann entlasten und beschleunigen, sie verspricht Übersicht und Effizienz. Zugleich greift sie tiefer in Entscheidungsprozesse ein, als es auf den ersten Blick scheint. Urteile entstehen nicht mehr allein im Moment bewusster redaktioneller Abwägung, sondern sind zunehmend vorgelegt. Welche Folgen das für journalistische Urteilkraft hat, wird in Kap. 3 ausführlicher diskutiert. Für den Journalismus ist das, im Ganzen gesehen, kein technisches Detail, sondern *ein Wendepunkt im Rahmen seiner Entstehungsgeschichte*. Denn wenn Systeme, die nichts wissen, beginnen, vieles vorzusondieren, steht nicht ihre Leistungsfähigkeit zur Debatte – sondern unser Urteilsvermögen. Die entscheidende Frage ist nicht, ob KI richtige Antworten liefern kann, sondern ob Redaktionen noch erkennen, wann und warum sie ihr widersprechen müssen.

Ein zentrales Missverständnis in der aktuellen KI-Debatte besteht unseres Erachtens darin, die Leistungsfähigkeit der Technologie als linearen Fortschritt zu denken. In dieser Logik wird unterstellt: Je besser KI wird, desto umfassender lässt sie sich einsetzen. Die empirische Realität deutet jedoch auf das Gegenteil hin. Aktuelle internationale Forschung

beschreibt die Fähigkeiten von KI-Systemen nicht als kontinuierliche Kurve, sondern als „gezackte Grenze“ („jagged frontier“, Abb. 5): Systeme zeigen in einzelnen Aufgabenfeldern hohe Kompetenz, versagen jedoch abrupt in anderen – insbesondere dort, wo Kontextualisierung, Abwägung oder situatives Urteil erforderlich sind. Dieses Muster ist kein Zufall, sondern systemisch bedingt, da Interface-Design, Trainingslogiken und Nutzungskontexte Fehlzuschreibungen von Kompetenz begünstigen.²¹

Der in der folgenden Abbildung 5 markierte Punkt (★) verdeutlicht die strukturelle Ambivalenz aktueller KI-Systeme: Sie liefern in klar umrissenen Aufgabenfeldern zuweilen sehr verlässliche Ergebnisse, brechen jedoch bei Kontextwechseln oder normativ offenen Fragestellungen abrupt ein.²² Der Klassiker: Für Journalismus und Wissensarbeit entsteht daraus ein Spannungsfeld zwischen realen Effizienzgewinnen und dem Risiko systematischer Überschätzung, wenn maschinelle Leistungsfähigkeit als universelle Kompetenz missverstanden wird.

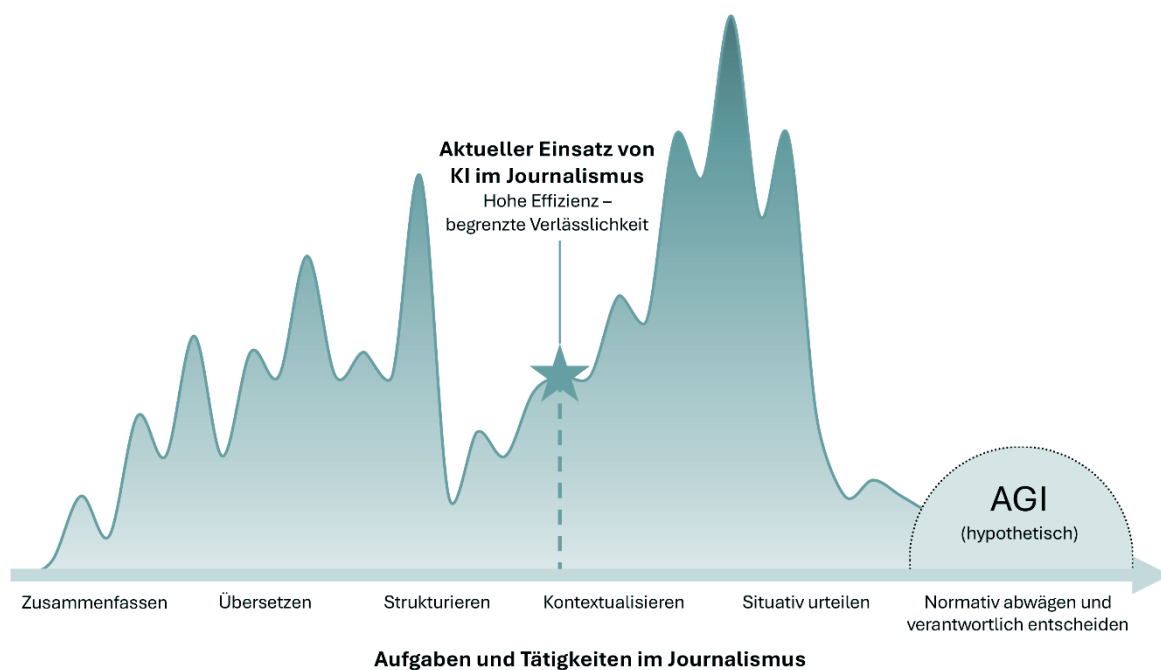


Abbildung 5: Die gezackte Grenze der KI-Leistungsfähigkeit im Hinblick auf Aufgaben und Tätigkeiten im Journalismus (eigene Darstellung nach Dell’Acqua und Kolleg:innen)²³

In Abbildung 5 steht die Markierung „AGI (hypothetisch)“ für ein oft beschworenes, aber derzeit rein theoretisches Leistungsniveau jenseits heutiger generativer Systeme. Wichtig für den Journalismus sind die heute bereits hoch automatisierten Aufgaben wie Zusammenfassen, Übersetzen und Strukturieren, während Kontextualisierung, situatives Urteil und normative Abwägung bislang klar menschlicher Verantwortung überlassen bleiben.

Abbildung 5 verdeutlicht, inwiefern KI innerhalb klar umrissener Aufgaben erstaunlich präzise arbeiten kann. Gleichzeitig zeigt sie erhebliche Schwächen in angrenzenden Bereichen, die für journalistische Arbeit zentral sind: Einordnung, Bewertung von Quellen, Plausibilitätsprüfung und ethische Abwägung. Plastisch gesprochen: Ein Sprachmodell kann binnen Sekunden juristische Schriftsätze verfassen, scheitert aber regelmäßig daran, ein einfaches Kinderspiel wie Schnick-Schnack-Schnuck strategisch konsistent über mehrere Runden zu spielen – weil es Kontext, Gedächtnis und situative Intentionalität nicht stabil zusammenführt.

Für Redaktionen ist diese „gezackte Grenze“ besonders relevant, weil journalistische Praxis nicht aus isolierten Einzelaufgaben besteht. Journalismus operiert gerade in den Übergangszonen: zwischen Fakt und Bedeutung, zwischen Geschwindigkeit und Sorgfalt, zwischen Information und Denkdziplin. Genau dort ist KI am anfälligsten – aber genau dort wird sie zunehmend eingesetzt.

Die eigentliche Gefahr liegt daher nicht in der Überforderung der Systeme, sondern in ihrer *situativen Überschätzung*. KI wirkt zuverlässig, solange sie sich innerhalb klar definierter Kompetenzbereiche bewegt. Sie signalisiert Sicherheit auch dort, wo sie lediglich statistisch plausibilisiert. Fehler sind dabei nicht immer offensichtlich, sondern werden oft sprachlich sauber und logisch anschlussfähig gemacht (*Affirmation-Bias*). Gerade das macht sie schwer erkennbar – insbesondere unter Zeitnot und Produktionsdruck.

Für den Journalismus folgt daraus eine zentrale Erkenntnis: KI ist weder grundsätzlich ungeeignet noch im umfassenden Sinn zuverlässig. Ihr Einsatz verlangt keine pauschale Zustimmung und kein kategorisches Misstrauen, sondern *Kontextsensibilität*. Entscheidend ist nicht, ob KI ‚gut genug‘ ist, sondern wo, wie und unter welchen Bedingungen sie eingesetzt wird. Genau hier entscheidet sich, ob Redaktionen ihre Kulturtechnik des Zweifels – also das Prüfen, Abwägen und Kontextualisieren – eher stärken oder schleichend aus der Hand geben.

Die „gezackte Grenze“ markiert deshalb keinen technischen Mangel, sondern sie ist eine Herausforderung für *organisationale Resilienz*. Sie zwingt Redaktionen dazu, Haftung und Rechenschaftspflichten nicht nur individuell, sondern strukturell zu verorten: durch klare Zuständigkeiten, transparente Prozesse und bewusste Grenzziehungen zwischen maschineller Unterstützung und menschlichem Urteil. KI-Resilienz bedeutet in diesem Zusammenhang nicht, Fehler zu vermeiden, sondern Systeme so zu gestalten, dass Fehler erkannt, korrigiert und verantwortet werden können.

2.3 Machtfragen: Plattformen, Modelle und neue Gatekeeper

Mit der Verbreitung generativer KI verändert sich die Machtbalance im digitalen Raum enorm – aber leiser und tiefgreifender als bei früheren Umbrüchen in der Plattformökonomie. Während Suchmaschinen und soziale Netzwerke seit Jahren darüber entscheiden, welche Inhalte sichtbar werden, tritt nun eine weitere Ebene hinzu:

KI-Systeme fungieren nicht mehr nur als Distributionsinfrastruktur, sondern zunehmend als vorstrukturierende Instanz von Öffentlichkeit.

Das heißt: Sie beantworten Fragen, fassen Informationen zusammen, priorisieren Themen – oft, ohne dass journalistische Quellen noch sichtbar werden. Die klassische Plattformlogik von Reichweite und Engagement wird ergänzt durch ein neues Prinzip der Antwortgebung: schnell, plausibel, autoritativ wirkend.

Diese Verschiebung ist für den Journalismus von besonderer Tragweite. In sogenannten „Zero-Click“-Umgebungen erhalten Nutzende Antworten, ohne journalistische Angebote überhaupt noch aufzurufen. Die Quelle verschwindet hinter der Antwort, der Kontext hinter der Zusammenfassung. Was früher ein bewusster Rezeptionsakt war – das Anklicken, Lesen, Einordnen –, wird zunehmend durch automatisierte Verdichtung via Chat-Funktion ersetzt. Damit verliert Journalismus nicht nur Reichweite, sondern Sichtbarkeit als kulturelle Praxis und als gesellschaftliche Institution ganz allgemein. Die Arbeit bleibt vorhanden, aber sie wird noch unsichtbarer gemacht.

Der Medienunternehmer und Analyst Burt Herman geht noch einen Schritt weiter und beschreibt diesen Kipppunkt als „People Zero“:²⁴ eine Medienwelt, in der kaum noch Menschen klassische journalistische Angebote direkt ansteuern, weil KI-Interfaces zur primären Oberfläche von Netzkommunikation geworden sind. Marken, Homepages und Reichweitenlogiken lösen sich im Dialog mit dem Chatbot auf; Journalismus erscheint dort, wenn überhaupt, als Fußnoten-Überbleibsel. Hermans Diagnose ist ernüchternd – und zugleich hoch anschlussfähig an ein resilienzorientiertes Verständnis von Medienwandel. Denn wenn Aufmerksamkeit nicht mehr über Klicks, sondern über Kontexte, Intimität und Anschlussfähigkeit suggeriert wird, verschiebt sich das journalistische Produkt fundamental.

Zukunftsfähiger Journalismus entsteht dann nicht (mehr) primär durch das bloße Bereitstellen von qualifizierten Inhalten, sondern durch Beziehung, konkreten Nutzen und die Fähigkeit, Menschen selbstwirksam zu machen – jenseits der Logik geometrischer KI-Antworten. Damit entstehen neue Gatekeeper, deren Einfluss subtiler, aber strukturell wirksamer ist als die der klassischen Plattformen. Nicht mehr Redaktionen – und nicht einmal mehr allein Reichweitenalgorithmen – entscheiden über Sichtbarkeit, sondern KI-

Modelle, deren Funktionsweise proprietär und für die Öffentlichkeit weitgehend opak bleibt. Diese Systeme aggregieren journalistische Inhalte, rekontextualisieren sie und formulieren sie neu, ohne an journalistische Kriterien wie Quellenkritik, Transparenz oder Verantwortlichkeit gebunden zu sein. Öffentlichkeit wird so nicht mehr nur verteilt, sondern synthetisch verfälscht – ein Strukturwandel, den der Journalismus bislang kaum systematisch würdigt.

Die Folge ist eine *strukturelle Asymmetrie*: Redaktionen tragen Verantwortung für Inhalte, während KI-Systeme selbst keine verantwortlichen Akteure sind; zwar lassen sich rechtliche Pflichten auf Hersteller, Anbieter oder Betreiber verteilen, doch die Systeme erzeugen Antworten ohne eine unmittelbar zurechenbare Instanz; sie wirken autoritativ, ohne selbst – außer vielleicht den Aktionären gegenüber – rechenschaftspflichtig zu sein.²⁵ Damit entsteht ein neues Spannungsfeld zwischen Effizienzgewinnen, ethischen Grauzonen und demokratischer Kontrolle. Journalismus bleibt also an normative Verpflichtungen gebunden – Transparenz, Korrekturfähigkeit, öffentliche Verantwortung –, während KI-Systeme außerhalb vergleichbarer institutioneller Haftungsräume operieren. Diese Entkopplung von Wirkung und Verantwortung ist ein demokratietheoretisches Problem.

Für Redaktionen bedeutet diese Entwicklung eine *doppelte ideologische Verschiebung ihrer Rolle und ihres ökonomischen Fundaments*:

Erstens verlagert sich die Wertschöpfung weiter weg vom publizierten Inhalt hin zur infrastrukturellen Ebene. Journalistische Arbeit wird zur Vorleistung für Systeme, die diese Inhalte aggregieren, verdichten und in neuer Form ausspielen – oft ohne Rückkopplung, Vergütung oder Sichtbarkeit.

Zweitens verlieren Redaktionen zunehmend die Kontrolle über den inhaltlichen Kontext, in dem ihre Inhalte erscheinen, interpretiert oder weiterverarbeitet werden. Artikel, Recherchen und Analysen zirkulieren nicht mehr als klar zuordenbare Beiträge, sondern als Teil maschineller Antwortketten, deren Logik nicht journalistisch, sondern algorithmisch getrieben ist.

Diese Entwicklung ist allem voran eine Frage *institutioneller Macht*. Wer KI-Modelle betreibt, Trainingsdaten kontrolliert und Distributionswege definiert, gestaltet die Bedingungen von Öffentlichkeit. Während Redaktionen an journalistische Sorgfaltspflichten, Pressekodizes und rechtliche Verantwortung gebunden sind, operieren KI-Systeme außerhalb vergleichbarer Rechenschaftsstrukturen.

Für den Journalismus verschärft sich damit ein allseits bekanntes Problem unter neuen Vorzeichen. Schon die Plattformökonomie setzte die redaktionelle Autonomie unter Druck, indem journalistische Eigenlogiken zunehmend von ökonomischen und

technologischen Bewertungsregimen überformt wurden.²⁶ Mit generativer KI wird diese Abhängigkeit nicht nur vertieft, sondern qualitativ verformt. Redaktionen geraten in die Position der nachgelagerten Akteure innerhalb von Infrastrukturen, die sie weder geschaffen haben noch beeinflussen können (anders als z.B. beim Zeitungsdruck oder TV-Sendestationen). Die Entscheidung darüber, welche Informationen relevant erscheinen, welche Perspektiven dominant werden und welche Stimmen marginalisiert bleiben, verschiebt sich von redaktionellen Prozessen hin zu ‚KI-fizierten‘ Architekturen: Öffentlichkeit wird zunehmend vorstrukturiert, bevor journalistische Arbeit überhaupt ansetzen kann.

Für das Verständnis von KI-Resilienz ist diese Machttektonik zentral. Denn Resilienz bezieht sich nicht allein auf den kompetenten Umgang mit neuen Werkzeugen, sondern auf die Fähigkeit von Organisationen, in solchen *asymmetrischen Infrastrukturen* handlungsfähig zu bleiben (vgl. Kap. 3): Sie fragt danach, wie redaktionelle Unabhängigkeit, Verantwortlichkeit und professionelle Standards unter Bedingungen gewahrt bleiben, in denen die Spielregeln von externen Playern definiert werden.

Ohne dieses Bewusstsein droht der Journalismus, schrittweise vom gestaltenden Akteur der Öffentlichkeit zum *Zulieferer einer von Plattformen und KI-Modellen dominierten Informationsordnung* zu werden.

2.4 Journalismus zwischen Effizienzgewinn und Identitätsverlust

Die Integration von KI in redaktionelle Prozesse verändert also nicht nur Arbeitsabläufe und Fragen der Urheberschaft, sondern das professionelle Selbstverständnis des Journalismus per se. Was zunächst als pragmatischer Effizienzgewinn erscheint – schnellere Recherche, automatisierte Zusammenfassungen, sprachliche Glättung –, entfaltet eine ambivalente Wirkung. Routinen werden beschleunigt, Produktionszyklen verkürzt, Entscheidungen vereinfacht. Gleichzeitig werden jene Elemente ausgehebelt, die journalistische Arbeit über Jahrzehnte ausgezeichnet und angetrieben haben: Zweifel, Akkuratess, und Originalität.

Dennoch übernehmen KI-Systeme Aufgaben, die lange als Kern journalistischer Kompetenz galten. Sie strukturieren Informationen, clustern Themen, generieren erste Textfassungen. Damit verschiebt sich der sichtbare Ort journalistischer Leistung: Sie liegt nicht mehr im veröffentlichten Produkt, sondern zunehmend im Hintergrund – in der Auswahl, Korrektur, Einordnung und Kontrolle maschineller Vorschläge. Journalistische Arbeit wird dadurch weniger greifbar, weniger erzählbar, weniger eindeutig zurechnungsfähig.

Zweifellos steigt die Effizienz, doch geht die privilegierte Domänenkompetenz des Journalismus verloren.

Für viele Journalist:innen geht damit ein Verlust an Selbstwirksamkeit einher. Wenn Maschinen Texte liefern, die formal korrekt, sprachlich souverän und inhaltlich plausibel erscheinen, stellt sich unausgesprochen die Frage nach der eigenen Rolle. Die Arbeit wird produktiver, aber nicht notwendigerweise selbstbestimmter. Der Beruf scheint leichter zu werden, verliert aber zugleich an Anerkennung – ein Wandel, der auch die professionelle Identität untergräbt.

Hinzu kommt ein strukturelles Problem der *Verantwortungsdiffusion*. Fehler, die früher eindeutig menschlichen Entscheidungen zugeordnet werden konnten, erscheinen nun als technische Abweichungen oder werden tendenziell als technische ‚Fehltritte‘ oder Einzelfälle verharmlost (vgl. prominente Fälle aus dem deutschen Onlinejournalismus, u. a. bei Burda, dem Ippen Verlag, bei „Spiegel Online“ oder dem ZDF).²⁷ Verantwortung verteilt sich zwischen Mensch und System – oder wird stillschweigend an das Werkzeug delegiert. Aussagen wie: ‚Das hat die KI gemacht‘ markieren eine neue Grauzone journalistischer Zuständigkeit. Sie entlasten kurzfristig, zerfressen aber langfristig die Grundlage journalistischer Glaubwürdigkeit: die Verantwortung, für eigene Entscheidungen einzustehen.

Diese Entwicklung wirkt unmittelbar in den redaktionellen Betrieb hinein. Wo Fehler früher als Ausdruck menschlicher Begrenztheit galten – und korrigiert werden konnten –, erscheinen sie heute als systembedingt Kollateralschäden. Die klassische Fehlerkultur, die Transparenz, Korrektur- und Lernfähigkeit miteinander verbindet, gerät unter Druck. Redaktionen laufen Gefahr, Haftungsfragen zu entpersonalisieren und damit genau jene Vertrauensressource zu schwächen, die sie im digitalen Raum am dringendsten benötigen. In vielen Medienhäusern zeigt sich diese zunehmende Unsicherheit deutlich: in vorsichtig formulierten Prompting-Regeln, abstrakten hausinternen Leitfäden und der latenten Angst, durch KI ersetzt zu werden, bevor man gelernt hat, mit ihr souverän zu arbeiten.

Auch hierauf strahlt der „ELIZA-Effekt“ und damit die ‚Psychologie der KI‘ ab: Die Orientierungslosigkeit darüber, was KI leisten kann – und was nicht –, wirkt tief in das redaktionelle Selbstverständnis hinein – etwa durch Affektregulation, Affirmation oder Autoritätshörigkeit.²⁸ Die leise, aber wirkmächtige Frage *Kann die KI das nicht besser?* begleitet bereits etliche redaktionelle Entscheidungen, ohne offen diskutiert zu werden.

Für den Journalismus entstehen dadurch tiefgreifende strukturelle Herausforderungen, die in den Folgekapiteln diskutiert werden. Die Einführung von KI berührt drei zentrale Grundfragen des Berufs:

1. **Verantwortung:** Wer trifft Entscheidungen?
2. **Haftung:** Wer steht für Inhalte ein?
3. **Qualität:** Woran misst sich journalistische Güte, wenn Produktionsprozesse teilautomatisiert sind?

Diese Fragen lassen sich weder durch bessere Tools noch durch zusätzliche Effizienz beantworten. Sie erfordern organisatorische Klarheit, institutionelle Regeln und eine bewusste Auseinandersetzung mit publizistischer Verantwortung – jenseits technischer Machbarkeit.

Kapitel 2 hat gezeigt, dass der Einsatz von KI im Journalismus weniger an fehlender Technologie scheitert als an offenen Fragen von Einbettung, Zuständigkeit und Verantwortung. KI greift in Entscheidungsprozesse, Rollenverständnisse und professionelle Selbstbilder ein. Diese Verschiebungen lassen sich nicht allein über technische Leistungsbeschreibungen angemessen erfassen. Vor diesem Hintergrund stellt sich daher die Frage, mit welchen begrifflichen und analytischen Werkzeugen sich beschreiben lässt, wie Journalismus unter diesen Bedingungen selbsterhaltend arbeiten kann und dabei verantwortungsvoll und orientierungsgebend bleibt. Genau hier setzt das folgende Kapitel an und stellt das Konzept der *KI-Resilienz* vor.

3. Resilienz als Schlüsselkonzept im Journalismus – Bewahren, Lernen, Anpassen

Worauf dieses Kapitel Antworten gibt:

1. Warum Resilienz im KI-Zeitalter als Antwort auf den strukturellen Umbruch im Journalismus nötig ist und weshalb technologische Lösungen allein nicht genügen.
2. Wie Resilienz begrifflich als Prozess von Bewahren, Lernen und Anpassen unter digitalen Bedingungen verstanden werden sollte.
3. Wie sich die drei Ebenen der Resilienz verbinden, wenn individuelle Urteilkraft, organisationale Praxis und gesellschaftliche Verantwortung zusammenwirken.
4. Wo sich im redaktionellen Alltag neue Spannungsfelder und Praxismuster im Umgang mit KI zeigen – und wann Resilienz tragfähig bleibt oder fragil wird.
5. Warum Resilienz als Haltung und Ressource sowohl professionelle Selbstvergewisserung im Journalismus verlangt als auch demokratische Stabilität stützt.

3.1 Warum Resilienz?

Vom KI-Bruch zur Digitalen Resilienz als Orientierungsrahmen

Der Einsatz von KI im Journalismus stellt weniger die technische Machbarkeit infrage als die journalistische Verantwortung.²⁹ Entscheidungen werden zunehmend durch Systeme vorbereitet, Inhalte vorstrukturiert und Relevanzen algorithmisch mitgeprägt, ohne dass stets klar erkennbar bleibt, wer dafür die publizistische Verantwortung trägt. Genau hier kommt die Frage für Medienschaffende nach einem resilienteren Umgang mit KI ins Spiel: Der Resilienzbegriff zielt nicht auf technische (z.B. leichter zu bewerkstelligende) Lösungen, sondern bietet einen ganzheitlichen analytischen Rahmen, um mit entstehender oder dauerhafter Unsicherheit so umzugehen, dass professionelle Autonomie und Selbstwirksamkeit erhalten bleiben.

Resilienz ist dabei freilich kein Konzept, das exklusiv auf die journalistische Praxis abzielt. In unterschiedlichen gesellschaftlichen Kontexten beschreibt der Begriff die Fähigkeit von Individuen, Organisationen oder ganzen Systemen, mit Störungen aller Art so umzugehen, so dass grundlegende Orientierungen erhalten bleiben. In der Psychologie verweist Resilienz auf den Umgang mit Belastungen, ohne daran zu zerbrechen.³⁰ In organisationalen Zusammenhängen steht Resilienz für Lern- und Anpassungsfähigkeit unter

Krisenbedingungen.³¹ Und in gesellschaftlichen Debatten bezeichnet Resilienz darüber hinaus die Widerstandskraft demokratischer Institutionen – oder Demokratie als Ganzem – gegenüber externen und internen Erosionen.³² Gemeinsam ist diesen Lesarten, dass Resilienz nicht Stabilität um jeden Preis meint, sondern die Fähigkeit, mit Veränderung umzugehen, ohne die eigene Orientierungs- und Gestaltungsfähigkeit zu verlieren – ebenso wenig wie die Urteilstkraft, also die Fähigkeit, unter dem Eindruck von Unsicherheit abzuwägen, Bedeutung herzustellen und Entscheidungen verantwortbar zu begründen.

Mit der fortschreitenden Digitalisierung hat sich dieses Verständnis weiter ausdifferenziert. **Digitale Resilienz** beschreibt den Umgang mit technischen Systemen, die dynamisch, intransparent und nur begrenzt kontrollierbar sind.³³ Sie entsteht dort, wo Abhängigkeiten anerkannt, aber nicht bedingungslos akzeptiert werden; wo Effizienzgewinne genutzt werden, ohne Verantwortung auszulagern; und wo Unsicherheit nicht verdrängt, sondern für individuelle Reflexion und organisationales Lernen genutzt wird.

Digitale Resilienz erkennt technologische Komplexität an, schützt vor dem Irrglauben, technologische Neuerungen ließen sich vollständig beherrschen oder neutralisieren und stärkt nachweislich die eigene Handlungsfähigkeit durch:

1. **Urteilskompetenz und professionelle Befähigung:** Digitale Resilienz entwickelt sich dort, wo journalistische Arbeit bewusst reflektiert und eingeübt wird. Entscheidend ist, dass Journalist:innen wie auch Publikum lernen, Einordnungen vorzunehmen, Quellen kritisch zu prüfen und Entscheidungen nachvollziehbar zu begründen. Weiterbildung, Coaching und praxisnahe Lernformate spielen dabei eine zentrale Rolle. Sie schaffen Räume, in denen redaktionelle Dilemmata sichtbar werden und gemeinsam bearbeitet werden können. Initiativen wie *Journalismus macht Schule*³⁴ verdeutlichen, wie journalistische Arbeitsweisen vermittelt werden können, indem reale Entscheidungsprozesse offengelegt und gerade mit jüngeren Mediennutzenden diskutiert werden.
2. **Geteilte Lernstrukturen und kooperative Praxis:** Resilienz entsteht nicht isoliert in einzelnen Redaktionen. Sie wächst dort, wo Medien mit Wissenschaft, Zivilgesellschaft und lokalen Akteuren zusammenarbeiten und Wissen austauschen. Netzwerke zur Überprüfung von Informationen oder zur gemeinsamen Analyse digitaler Kommunikationsmuster tragen dazu bei, neue Formen von Irreführung frühzeitig zu erkennen und öffentlich einzuordnen. Solche Kooperationen stärken den öffentlichen Informationsraum, weil sie Verantwortung verteilen und kollektive Lernprozesse ermöglichen.

3. **Gesellschaftliche Handlungsfähigkeit gegenüber Plattformdynamiken:** Eine zentrale Herausforderung liegt in der wachsenden Bedeutung sozialer Netzwerke für öffentliche Debatten. Reichweitenmechaniken, Empfehlungslogiken und KI-gestützte Inhalte verändern, welche Themen sichtbar werden und wie Diskussionen verlaufen. Digitale Resilienz bedeutet hier, diese Dynamiken politisch und gesellschaftlich zu adressieren. Initiativen wie *Save Social*³⁵ zeigen, wie sich medienpolitische Forderungen, öffentliche Aufklärung und journalistische Perspektiven verbinden lassen, um demokratische Kommunikationsräume langfristig zu sichern.
4. **Vorbereitete Reaktionsfähigkeit in angespannten Lagen:** Resilienz zeigt sich im Alltag besonders dann, wenn öffentliche Akteur:innen souverän bleiben, obwohl sich Gerüchte, manipulierte Inhalte oder irreführende Darstellungen schnell verbreiten. Hilfreich sind vorab geklärte Zuständigkeiten, eingespielte Kommunikationswege und verständliche Erklärangebote, die im Bedarfsfall aktiviert werden können. So wird es möglich, zeitnah zu erläutern, was bekannt ist, was geprüft wird und wie Sachverhalte einzuordnen sind. Diese Form der Vorbereitung im Sinne von ‚expect the unexpected‘ schafft Vertrauen, weil sie Transparenz ermöglicht, ohne vorschnelle Gewissheiten zu behaupten.

3.2 Resilienz begrifflich klären: Bewahren, Lernen, Anpassen

Resilienz wird in der Forschung als dynamische Bewegung verstanden. Im Kern beschreibt sie die Fähigkeit von Akteuren oder Systemen, auf Störungen oder Wendepunkte (im Arbeitsalltag) so zu reagieren, dass grundlegende Funktionen und Orientierungen erhalten bleiben. Diese Prozesshaftigkeit ist zentral, um Resilienz analytisch von bloßer mentaler Robustheit oder kurzfristigen Anpassungsleistungen zu unterscheiden. Resiliente Systeme zeichnen sich nicht dadurch aus, dass sie unveränderlich bleiben, sondern dadurch, dass sie Veränderung – man könnte sagen: Werte – aktiv leben können, ohne ihre handlungsleitenden Prinzipien und institutionellen Bezugspunkte aus den Augen zu verlieren.

In der Resilienzforschung hat sich hierfür eine dreigliedrige Perspektive etabliert, die Bewahren, Lernen und Anpassen zusammendenkt:

1. **Bewahren** meint dabei keine Verweigerung gegenüber Veränderung, sondern die Fähigkeit, auch in Krisensituationen fundamentale Prinzipien zu verteidigen. Resiliente Akteure geben nicht jede normative Orientierung auf, nur weil äußere Bedingungen sich ändern – auch nicht im Extremfall. Gerade unter Druck zeigt sich, welche

journalistischen Werte und Standards als unverzichtbar gelten, welche verhandelbar sind und insofern bewahrt werden müssen.

2. **Lernen** bezeichnet die Fähigkeit, Störungen nicht ausschließlich als Ausnahmen zu behandeln, sondern zu versuchen, sie grundlegend zu verstehen und Lehren daraus zu ziehen. Resilienz setzt also voraus, dass Fehler, Irritationen und Brüche nicht individualisiert oder verdrängt werden, sondern in Lernprozesse überführt werden können. Eine lernende Organisation ist eine, die beispielsweise reflektiert, warum bestimmte organisationale Routinen an ihre Grenzen stoßen und welche bisherigen Praxisabläufe sich als unzureichend erweisen.
3. **Anpassen** schließlich beschreibt die Fähigkeit, Strukturen, Prozesse oder Rollen gezielt so zu verändern, dass sie auch unter neuen Bedingungen funktionieren und tragfähig bleiben. Wenn ein Unternehmen etwa KI-gestützte Software für Projektplanung oder Personalentscheidungen nutzt, lernen Mitarbeitende, mit neuen Oberflächen, Empfehlungen und Automatisierungen umzugehen. Resilient ist diese Anpassung jedoch nur dann, wenn die Verantwortung für Entscheidungen nicht an das System delegiert wird. Die Werkzeuge verändern, wie gearbeitet wird – nicht, wer entscheidet, priorisiert oder im Konfliktfall haftet.

Diese drei Dimensionen lassen sich analytisch klar voneinander trennen, sind praktisch jedoch eng miteinander verschränkt. Resilienz entsteht nicht dadurch, dass Journalist:innen in einzelnen Situationen ihre professionellen Standards bewahren, lernen oder sich anpassen, sondern diese drei Ebenen als Zusammenspiel praktizieren. Systeme, die ausschließlich bewahren, gelten als rückwärtsgewandt; solche, die ihre Methoden ständig anpassen, verlieren die professionelle Gravitas; und eine Organisation, die nicht lernbereit ist, kann ihre Erfahrungen nicht in handlungsleitendes Wissen übersetzen.

Für digitale Kontexte ist diese begriffliche Schärfung besonders relevant, da hier Störungen alltäglich sind. Nicht etwa im Sinne von Defekten, sondern von Irritationen: Mit ‚Störungen‘ sind Eingriffe in gewohnte Arbeits- und Entscheidungsabläufe gemeint. Neue Technologien wirken in Organisationen hinein, indem sie Routinen verändern, Abläufe beschleunigen und neue Abhängigkeiten erzeugen. Dadurch entstehen Situationen, in denen Zuständigkeiten neu ausgehandelt werden müssen oder gewohnte Bewertungsmaßstäbe nicht mehr greifen.

Was besonders schwer wiegt: *Digitale Technologien erzeugen Störungen nicht punktuell, sondern dauerhaft.* Resilienz kann hier nicht als Rückkehr zum vorherigen Zustand verstanden werden (‚Stehaufmännchen-Prinzip‘), sondern als fortlaufende Aushandlung

unter Bedingungen permanenter Veränderung. Digitale Resilienz bezeichnet mithin die Fähigkeit, mit Unsicherheiten und technologischen Abhängigkeiten so umzugehen, dass die Souveränität im Digitalen eher noch gestärkt wird.

3.3 Drei Ebenen journalistischer Resilienz: Individuum, Organisation, Gesellschaft

Resilienz entfaltet ihre analytische Stärke erst dann, wenn sie nicht auf eine einzelne Handlungsebene verkürzt wird. Für den Journalismus ist es entscheidend, Resilienz als *Mehrebenenkonzept* (Abb. 6) zu begreifen: Individuelle, organisationale und gesellschaftliche Dimensionen sind dabei analytisch zwar unterscheidbar, in praxi aber eng miteinander verschränkt und schwerlich trennbar. Auch lassen sich Defizite auf einer Ebene nicht dauerhaft durch die Stärken auf einer anderen Ebene kompensieren.

1. Auf der *individuellen Ebene* gehört zum „resilienten Journalismus“³⁶, eigene Entscheidungen sicher zu treffen, sie kritisch zu reflektieren und Verantwortung als integralen Bestandteil professionellen Handelns wahrzunehmen. Urteilstkraft bleibt dabei an die handelnde Person gebunden, auch unter Bedingungen von Zeitdruck, Unsicherheit und widersprüchlichen Anforderungen. Individuelle Resilienz äußert sich in Bezug auf die digitale Transformation daher nicht in technischer Versiertheit allein, sondern in der Verteidigung übergreifender journalistischer Autonomie gegenüber Einflüssen von jedweder Seite. Sie ist auch keine rein persönlich-psychische Eigenschaft, sondern immer eingebettet in organisationale Rahmenbedingungen.
2. Die *organisationale Ebene* bildet daher den *zentralen Hebel journalistischer Resilienz*. Hier entscheidet sich, ob individuelle Resilienz wirksam werden kann oder ins Leere läuft. Organisationale Resilienz zeigt sich dort, wo Redaktionen den Einsatz digitaler und KI-gestützter Werkzeuge in konkrete Arbeitsabläufe übersetzen: etwa indem festgelegt ist, wann KI bei Recherche, Zusammenfassungen oder Textentwürfen eingesetzt werden darf, wer Ergebnisse prüft und wer bei Unsicherheiten die Entscheidung übernimmt. Fehlen solche Klärungen, entstehen Grauzonen, in denen Verantwortung faktisch individualisiert wird. Klare Zuständigkeiten bei KI-gestützten Beiträgen, nachvollziehbare Prüfungen vor Veröffentlichung und feste Anlaufstellen, wenn Zweifel an Qualität, Tonalität oder Fakten entstehen, können hier helfen. Ebenso wichtig ist der Umgang mit Fehlern: Wo KI-unterstützte ‚Pannen‘ offen besprochen und nicht einzelnen Personen zugeschrieben werden, sind Redaktionen lernfähig statt verunsichert. Dies schützt professionelle Standards vor einer schleichenden Erosion.

3. Auf der *gesellschaftlichen Ebene* schließt sich der Kreis. Hier verweist journalistische Resilienz auf die Rolle des Journalismus als Infrastruktur demokratischer Öffentlichkeit. Resiliente Medien tragen dazu bei, Vertrauen, Transparenz und Verständigungsfähigkeit auch unter digitalen Vorzeichen zu sichern. Diese Ebene wird oft erst dann sichtbar, wenn sie erodiert: wenn sich Desinformation ungehindert ausbreitet, wenn journalistische Leistungen für Nutzende nur noch schwerlich erkennbar sind oder wenn öffentliche Debatten zum Selbstzweck einer Aufmerksamkeitsökonomie ohne gemeinsame Bezugsbasis avancieren. Gesellschaftliche Resilienz hängt daher wesentlich davon ab, ob Journalismus seine orientierende Funktion auch unter Bedingungen algorithmischer Vorstrukturierung behaupten kann.³⁷

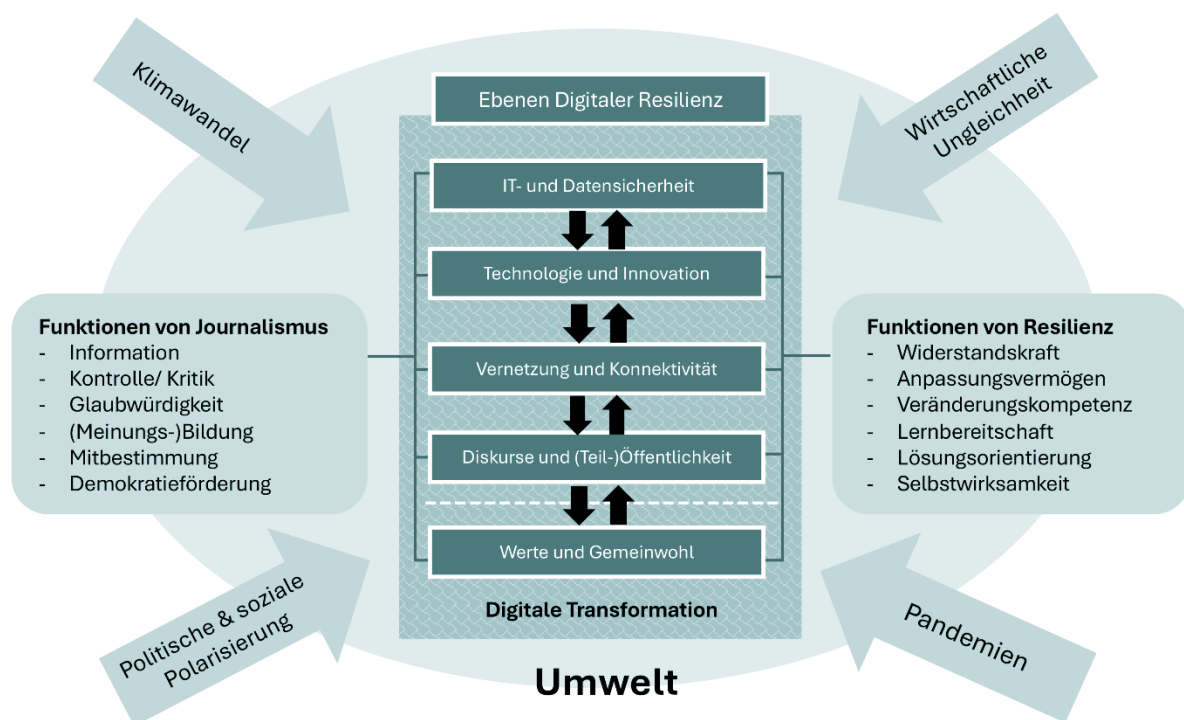


Abbildung 6: Herausforderungen Digitaler Resilienz als Mehrebenen-Modell (eigene Darstellung auf Basis von Merkel³⁸, Atteneder u.a.³⁹ sowie Hepp/Hasebrink⁴⁰)

Entscheidend ist damit die Verschränkung dieser Ebenen: Individuelle Resilienz ohne organisationale Unterstützung bleibt fragil. Organisationale Resilienz ohne gesellschaftliche Einbettung droht, sich in Effizienzlogiken zu verengen. Und gesellschaftliche Erwartungen an einen ‚resilienten Journalismus‘ lassen sich nicht erfüllen, wenn individuelle und organisationale Voraussetzungen fehlen.

Das Konzept macht sichtbar, dass technologische Umbrüche nicht isoliert auf der Ebene einzelner Tools wirken, sondern professionelle Rollen, organisationale (Redaktion-)Kulturen und gesellschaftliche Erwartungen zugleich tangieren. Dieses ‚Hinein- und Hinauszoomen‘ ist zentral, um KI nicht nur als Effizienzfrage, sondern als grundlegende Strukturveränderung journalistischer Praxis in actu zu begreifen. Resilienz im Journalismus ist daher ein *relationales* Konzept. Diese Mehrebenen-Perspektive ermöglicht es, digitale Transformation und KI nicht als isolierte Herausforderung zu behandeln, sondern als Prüfstein für bestehende Strukturen journalistischer Praxis.

3.4 Resilienz im Alltag:

Spannungsfelder, Routinen und organisationale Praxis

Die Stärke journalistischer Resilienz zeigt sich rund um die Uhr im redaktionellen Alltag: Dort, wo unter Zeitdruck entschieden wird, immer neue Schnittstellen zu technischen Systemen und Plattformen integriert werden sollen und *Verantwortung* situativ ausgehandelt werden muss zeigt sich, ob Resilienz mehr ist als ein normativer Anspruch. Digitale Transformation – und aktuell insbesondere KI – wirkt in diesem Kontext weniger als singulärer Bruch denn als Beschleuniger bestehender Routinen, Machtverhältnisse und kultureller Selbstverständlichkeiten in Redaktionen.

An diesem Punkt wird sichtbar, dass Resilienz im Journalismus auf sehr unterschiedliche Weise entstehen kann. Redaktionen und Medienhäuser begegnen vergleichbaren Herausforderungen mit je eigenen Verfahrensweisen: Manche setzen auf formale Regeln, interne Leitbilder und Governance-Strukturen. Andere investieren vorrangig in Kompetenzaufbau, Weiterbildung und interne Lernprozesse. Wieder andere orientieren sich stärker an ökonomischer Effizienzsteigerung, reagieren vorsichtig-abwägend oder wählen bewusst experimentelle, öffentlichkeitswirksame Formen und Formate des Umgangs mit dem KI-Wandel.

Diese unterschiedlichen Herangehensweisen lassen sich nicht auf einzelne Entscheidungen oder kurzfristige Strategien zurückführen. Sie sind Ausdruck langfristig gewachsener institutioneller Konstellationen, organisationaler Routinen und professioneller Rollenselbstverständnisse. Je nach Funktion eines Medienunternehmens im Mediensystem und auf dem Medienmarkt entwickeln Redaktionen unterschiedliche Umgangsweisen mit Veränderungsdruck, Verantwortung und Unsicherheit. Um diese unterschiedlichen Praxismuster systematisch vergleichbar zu machen, ordnet Abbildung 7 die beobachtbaren Formen journalistischer Resilienz entlang zweier grundlegender Dimensionen; sie

dient als Orientierungsrahmen für die anschließenden Fallbeschreibungen und Bewertungen:

Auf der vertikalen Achse ist kategorisiert, in welchem Maße Redaktionen die KI aktiv gestalten oder überwiegend reaktiv einsetzen. Die horizontale Achse beschreibt den Grad redaktioneller Souveränität, also die Frage, wie stark journalistische Entscheidungen, Verantwortung und Deutungshoheit unter KI-Bedingungen in der Redaktion selbst verbleiben. Aus der Kombination beider Dimensionen ergeben sich unterschiedliche Formen von KI-Resilienz, die nicht als klar abgrenzbares Stufenmodell zu verstehen sind, sondern als *idealtypische Orientierungspunkte*: Sie markieren, wie Organisationen mit ähnlichen technologischen Voraussetzungen sehr unterschiedliche Wege einschlagen – von instrumenteller Nutzung unter hohem Effizienzdruck bis hin zu kompetenz- oder regelbasierten Ansätzen, die auf langfristige Stabilität und nachhaltige Selbstbestimmung zielen. Die Abbildung soll damit als analytisches Hilfsmittel dienen, um die im Folgenden beschriebenen Praxisbeispiele vergleichbar zu machen und ihre jeweiligen Spannungsfelder, Stärken und Grenzen zu bewerten.

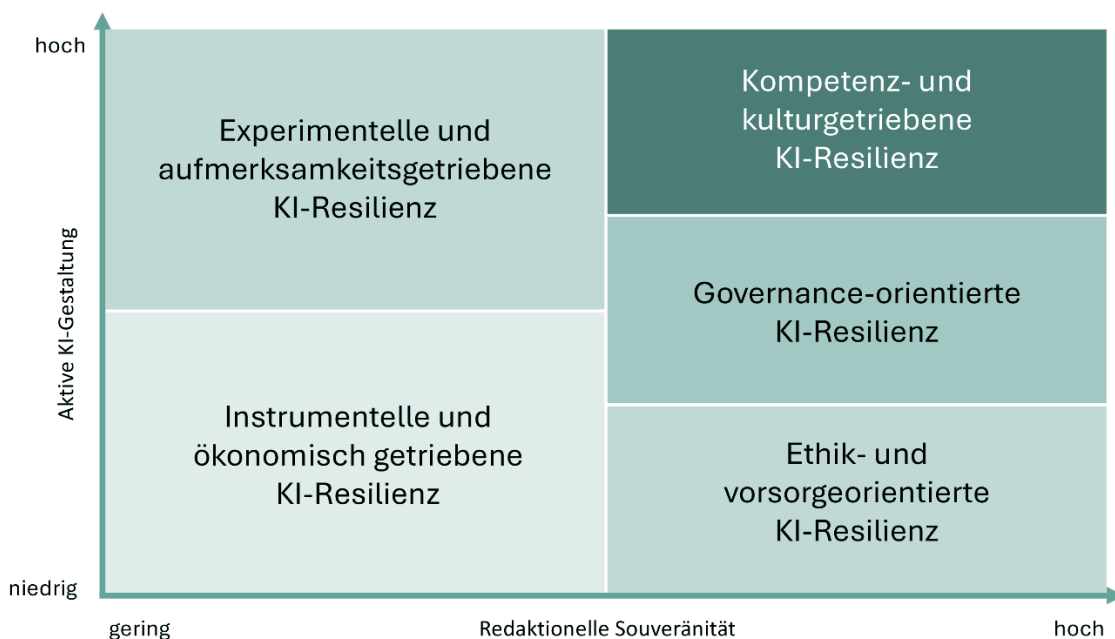


Abbildung 7: Typologie journalistischer KI-Resilienz entlang von redaktioneller Souveränität und aktiver KI-Gestaltung (eigene Darstellung)

A. Governance-orientierte KI-Resilienz:

Institutionelle Selbstbindung unter Innovationsdruck

Governance-orientierte KI-Resilienz entsteht dort, wo Organisationen den Einsatz von KI nicht primär als technische Frage behandeln, sondern als institutionelle Herausforderung, die *klare Zuständigkeiten, Begrenzungen und Verfahren* erfordert. In diesem Verständnis wird Resilienz weniger durch Geschwindigkeit oder Experimentierfreude erzeugt als durch *bewusste Selbstbindung*. KI wird organisatorisch erst gerahmt, bevor sie operativ skaliert wird. Dieses Muster ist insbesondere in Redaktionen zu beobachten, deren Legitimität in hohem Maße auf Vertrauen, Verlässlichkeit und öffentlicher Verantwortung beruht.

- Beispielsweise beim „Bayerischen Rundfunk“ (BR) zeigt sich diese Form der Resilienz in der konsequenten Einbettung von KI in bestehende journalistische Normen und Arbeitsweisen. Der Einsatz von KI wird als eigenständige Innovationsagenda verfolgt, allerdings in Form einer konsequenten Erweiterung redaktioneller Infrastruktur. Zentral ist dabei die klare Abgrenzung: KI darf unterstützen, aber keine Entscheidungsmacht übernehmen. Uli Köppen, Leiterin des AI + Automation Lab beim „Bayerischen Rundfunk“, betont, dass KI im BR nicht als autonome Produktionsinstanz eingesetzt wird, sondern durch interdisziplinäre Teams, interne Workshops und klare redaktionelle Leitlinien in journalistische Arbeitsprozesse eingebettet – etwa wenn KI-Tools assistierende Aufgaben übernehmen, gleichzeitig aber kontinuierlich erklärt, geprüft und auf ihre Integrationsfähigkeit in bestehende Strukturen untersucht werden: „Deshalb haben wir intensiv an externen und internen Guidelines für unsere Mitarbeitenden gearbeitet, [...] damit sie wissen, wie sie ein Tool wie ChatGPT benutzen, welche Daten sie da bedenkenlos reingeben können“, erklärt Köppen im Hintergrundgespräch. Resilienz entsteht hier also durch mehr Orientierung und Verlässlichkeit, weniger durch schnelle Anpassung.⁴¹

www.br.de

- Mit Blick ins europäische Ausland zeigt sich ein ähnliches, aber anders akzentuiertes Muster. Bei der „British Broadcasting Corporation“ (BBC) wird KI in redaktionelle Abläufe über Pilotprojekte und Human-in-the-Loop-Strukturen kontrolliert eingeführt. Die BBC verbindet dadurch institutionelle Selbstbindung mit gezielter Erprobung. Governance dient nicht allein der Begrenzung, sondern der Legitimation von Innovation. Indem Experimente klar als solche ausgewiesen und organisatorisch flankiert werden, bleibt die redaktionelle Autorität sichtbar. Resilienz entsteht hier aus der Fähigkeit, Innovation zuzulassen, ohne die eigene Rolle als öffentliche Vertrauensinstitution zu unterminieren.⁴²

www.bbc.co.uk

- Das Beispiel der US-amerikanischen Nachrichtenagentur „Associated Press“ (AP) wiederum zeigt, wie ein global wirkendes Medienunternehmen seine Funktion als Referenzquelle für Medien aus aller Welt mit diversen kulturellen Hintergründen agiert, wenn es neue Technologien in seine Arbeitsprozesse und Produktionsabläufe integriert. Der Umgang mit generativer KI fällt hier auffällig restriktiv aus: KI wird vor allem als internes Werkzeug eingesetzt, etwa zur Unterstützung von Recherche, Strukturierung oder Übersetzung. Diese strategische Selbstbegrenzung schützt die Glaubwürdigkeit der Agentur und sichert ihre Rolle innerhalb des medialen Ökosystems.⁴³

www.ap.com

Gemeinsam ist diesen drei Beispielen, dass KI-Resilienz hier unter der Dominanz stark kontrollierter organisationaler Entscheidungen entsteht. ‚AI Governance‘ fungiert dabei zur Festigung professioneller Identität und entsprechender Alleinstellungsmerkmale im medialen Wettbewerb. Unterschiede zeigen sich vor allem in der Art, wie KI organisatorisch eingeeht wird: Der „Bayerische Rundfunk“ setzt auf klare normative Leitplanken, die vor dem operativen Einsatz greifen. Bevor KI breit genutzt wird, werden Zuständigkeiten, Grenzen und Freigabeprozesse festgelegt. Das schafft Orientierung im Alltag, verlangsamt jedoch die Einführung neuer Anwendungen. Die BBC verbindet Kontrolle damit, gezielt zu experimentieren. Neue KI-Anwendungen werden in Pilotprojekten erprobt, eng begleitet und regelmäßig überprüft. Innovation ist möglich, bleibt jedoch an Verantwortungsstrukturen und menschliche Letztverantwortung gebunden. „Associated Press“ verfolgt eine Strategie bewusster Selbstbegrenzung. Der Einsatz von KI bleibt in der öffentlichen Berichterstattung eingeschränkt und dient allenfalls als Hilfsmittel, um abgegrenzte Arbeitsprozesse effizienter zu organisieren.

In allen Mini-Case-Studies wird KI maßgeblich als unterstützendes Werkzeug behandelt. Diese Governance-orientierte Form der Resilienz ist stabil, erzeugt jedoch Spannungen: Sie schützt Vertrauen und Verantwortung, begrenzt aber potenziell Innovationsgeschwindigkeit und die Reichweite neuer Anwendungen.

B. Kompetenz- und kulturgetriebene KI-Resilienz: Lernen, Steuerung und strategische Selbstermächtigung

Diese Form journalistischer KI-Resilienz entsteht vor allem durch den geregelten Aufbau von Wissen und eine Redaktionskultur, die integrale Lernprozesse und reflexive Steuerung ermöglicht. Der Umgang mit KI wird nicht allein über Regeln oder Vorgaben gesteuert, sondern über ein gemeinsames Verständnis, routinierten Austausch und gezielte Weiterqualifizierung innerhalb der Redaktionen. Medienhäuser verfolgen damit vor allem das Ziel, neue Technologien aktiv zu verstehen und sinnvoll in ihre Arbeitsweisen einzubinden.

Dazu gehört, dass Mitarbeitende die Funktionsweise von KI kennen, ihre Ergebnisse einordnen können und ihren Einsatz an redaktionellen und strategischen Zielen ausrichten. Investitionen in Weiterbildung, spezialisierte Expertise und interne Lernstrukturen werden damit zur zentralen Voraussetzung, um potenziellen Abhängigkeiten vorzubeugen und technologische Irritationen bewusst auszugestalten:

- Die „Financial Times“ wird in diesem Zusammenhang häufig als paradigmatisches Beispiel angeführt. KI wird dort nicht als isoliertes Innovationsthema behandelt, sondern als Bestandteil redaktioneller Grundkompetenz.⁴⁴ Klare AI-Governance, strukturierte interne Trainings, interdisziplinäre Austauschformate, externe Weiterbildung als Beratungsinstrument sowie eigene Tools zielen darauf ab, ein gemeinsames Verständnis von Entwicklungspotenzialen und Grenzen zu entwickeln. Mithilfe eines „AI Playgrounds“ trainieren Redaktionsmitglieder KI-Anwendungen auf Grundlage eigener journalistischer Inhalte.⁴⁵ Entscheidend ist dabei weniger der konkrete Einsatz einzelner Anwendungen als die nach außen hin abgesicherte kulturelle Signalwirkung: KI soll die journalistische Arbeit verbessern, nicht ersetzen.

www.ft.com

- Ein verwandtes, aber stärker institutionell ausdifferenziertes Muster lässt sich bei der „New York Times“ beobachten: Hier wird KI-Resilienz über den gezielten Aufbau eigener Kapazitäten organisiert, etwa durch das mittlerweile achtköpfige „AI Team“, das redaktionelle, technische und strategische Denkweisen zusammenführt.⁴⁶ KI wird insgesamt nicht als generische Lösung betrachtet, sondern gezielt an bestehende Stärken der Medienmarke gekoppelt, etwa an investigative Recherche und Datenjournalismus. Damit stärkt die Redaktion ihre Position gegenüber Plattformen und externen Playern, indem sie ihre publizistischen Kernkompetenzen bündelt und strukturell-operative Zuständigkeiten in der Redaktion klar verortet.⁴⁷

www.nytimes.com

- „Zetland“ aus Dänemark verschiebt dieses Muster konsequent auf die Ebene redaktioneller Kompetenz und Organisationskultur. KI wird hier nicht primär über Prinzipien oder zentrale Steuerung eingeführt, sondern als selbstverständliches, eher randständiges Arbeitsmittel, dessen verantwortlicher Einsatz durch Lernen, Transparenz und interne Befähigung abgesichert wird. „Zetlands“ CEO Tav Klitgaard betont, dass KI dort zum Wettbewerbsvorteil eingesetzt werde, wo sie als selbstverständlicher Bestandteil redaktioneller Routinen eingesetzt wird. Wegweisend sei weniger die Fortexistenz formaler Regeln als ein geteiltes professionelles Verständnis darüber, wie und wofür KI genutzt wird – vergleichbar mit früheren digitalen Werkzeugen, deren Qualität sich erst im alltäglichen Gebrauch bewährt habe.⁴⁸ Entsprechend investiert „Zetland“ weniger in formale KI-Policys als in kontinuierliche Weiterbildung, offenen Erfahrungsaustausch und Entwicklung eigener redaktioneller KI-Hilfsmittel wie bei der Weiterentwicklung des Transkriptions-Tools „Good Tape“. Resilienz entsteht hier durch die Fähigkeit der Redaktion, technologische Werkzeuge kritisch zu verstehen, situativ einzusetzen und kollektiv weiterzuentwickeln. KI stärkt damit journalistische Autonomie, weil sie Kompetenz voraussetzt und kultiviert – nicht, weil sie zentral kontrolliert wird.

www.zetland.dk

Gemeinsam ist allen Beispielen, dass *KI-Resilienz* nicht defensiv verstanden wird. Sie entsteht hier nicht aus Zurückhaltung, sondern aus *aktiv-operativer Aneignung* mit journalistischer Verantwortung. Kompetenz- und kulturgetriebene Ansätze setzen voraus, dass Redaktionen bereit sind, ihre Zeit, Ressourcen und Aufmerksamkeit zu investieren. Diese Form der Resilienz ist dynamisch und anpassungsfähig, zugleich aber voraussetzungs-voll. Sie erfordert stabile Organisationsstrukturen, langfristige Strategien und eine redaktionelle Kultur, die Lernen als kontinuierlichen Prozess implementiert. *Wo diese Voraussetzungen gegeben sind, kann KI bestehende journalistische Souveränität und Autonomie stärken oder sogar neue Geschäftsbereiche entwickeln und mehr Reichweite zu erschließen helfen.*

C. Instrumentelle und ökonomisch getriebene KI-Resilienz:

Effizienz, Kostendruck und fragile Stabilisierung

Ein drittes Praxismuster journalistischer KI-Resilienz entsteht dort, wo der Einsatz von KI primär aus *ökonomischem Druck* heraus motiviert ist. In diesen Fällen wird KI weniger als organisationale oder institutionelle Herausforderung verstanden, sondern als Instrument zur Effizienzsteigerung, Kostensenkung oder Reichweitensicherung. Resilienz greift hier vor allem situativ-punktuell: Sie soll kurzfristig Stabilität herstellen, ohne dass dafür

tragfähige redaktionelle Strukturen aufgebaut werden. Genau darin liegt allerdings auch ihre Sollbruchstelle.

- Beim „USA Today Network“ (Gannett) wird seit Sommer 2023 generative KI erstmals in größerem Umfang direkt in der Content-Produktion eingesetzt, insbesondere in der automatisierten lokalen Sportberichterstattung. Die KI-Texte erschienen als reguläre redaktionelle Beiträge und waren für das Publikum unmittelbar sichtbar, ohne systematische redaktionelle Prüfung oder klare Kennzeichnung. Konkret kamen KI-generierte Kurzberichte zu High-School-Sportereignissen zum Einsatz, die automatisiert aus Ergebnisdaten erstellt wurden.⁴⁹ Ziel war eine schnelle Skalierung lokaler Inhalte bei begrenzten personellen Ressourcen. Organisatorisch blieb dieser Einsatz nur schwach abgesichert: Die Texte erschienen ohne ausreichende redaktionelle Prüfung oder entsprechende Kennzeichnung.⁵⁰ Als Nutzende des Angebots wiederholt auf inhaltliche Fehler, monotone Sprache und fehlenden Kontext hinwiesen, wurde das Projekt nach öffentlicher Kritik schließlich gestoppt.⁵¹ Der Umgang mit den offensichtlichen Problemen erfolgte reaktiv. Erst nach dem Rückzug wurden Qualitätsfragen, Kontrollmechanismen und Zuständigkeiten thematisiert. Der Fall zeigt exemplarisch, dass Effizienzgewinne durch KI ohne begleitende organisationale Strukturen schnell in Reputationsrisiken umschlagen können.⁵²

www.usatoday.com

- „Axios“ repräsentiert eine professionalisierte Variante instrumentell-ökonomisch getriebener KI-Resilienz. Der Einsatz von KI ist hier weniger durch akuten Krisendruck motiviert als durch eine konsequente wachstums- und skalierungsorientierte Unternehmensstrategie.⁵³ Mit der 2025 bekanntgegebenen Partnerschaft mit OpenAI, die sowohl den Ausbau von „Axios Local“ in weiteren Städten als auch den Zugriff auf KI-Technologien für interne Prozesse umfasst, wird KI explizit als infrastruktureller Hebel für Distribution, Produktentwicklung und Monetarisierung positioniert.⁵⁴ Resilienz wird hier zuallererst über wirtschaftliche Stabilisierung und Reichweitenlogik organisiert. „Axios“ setzt bei KI auf strategische Partnerschaften und effiziente Skalierung – weniger auf tief verankerte organisationaler Lernprozesse oder institutionelle Selbstbindung. Ansonsten hat sich die Redaktionsleitung klar dafür ausgesprochen, auf menschengemachten Journalismus zu setzen – über Clubmitgliedschaften, Live-Events, Newsletter, Mitgliederprogramme, den persönlichen Draht zur Redaktion und zum journalistischen Nachwuchs.⁵⁵

www.axios.com

- Auch die „Welt“ (Axel Springer) zeigt eine stark ausgeprägte instrumentelle und ökonomisch getriebene KI-Resilienz, allerdings in einer stärker strategisch gerahmten Variante. Ziel ist eine schrittweise Normalisierung von KI als redaktioneller Infrastruktur. KI wird dabei zunehmend nicht als optionales Werkzeug oder punktuelles Experiment behandelt, sondern als integraler Bestandteil journalistischer Produktionsprozesse, dessen Nutzung faktisch erwartet wird.⁵⁶ Der Einsatz zielt vor allem auf Effizienzgewinne, Skalierbarkeit und Geschwindigkeit in einem hochkompetitiven Marktumfeld. KI wird in der Redaktion gezielt eingesetzt – etwa durch automatisierte Textformate, KI-gestützte Recherchehilfen oder Moderationsformen (Avatare) im Bewegtbild.⁵⁷ In den vergangenen Jahren ist dieser KI-Einsatz öffentlich sichtbar gemacht und organisatorisch stärker verankert worden, etwa durch eigene Richtlinien und spezialisierte Führungsrollen („stellvertretender Chefredakteur für Künstliche Intelligenz“).⁵⁸ Die redaktionelle Verantwortung verbleibt formal bei den Redakteur:innen, verschiebt sich im Alltag jedoch in ein durch KI vorstrukturiertes Arbeitsregime. Resilienz entsteht in diesem Modell primär durch ökonomisch-technologische Anpassungsfähigkeit und operative Leistungsfähigkeit, nicht durch institutionelle Selbstbegrenzung oder ausgebaute Governance-Strukturen.

Vor diesem Hintergrund gewinnen auch frühere Rationalisierungsmaßnahmen im Springer-Konzern an Bedeutung: Der bereits 2023 angekündigte Stellenabbau ging der späteren strategischen Zuspitzung der KI-Nutzung voraus und verweist auf eine längerfristige ökonomische Restrukturierung, in deren Rahmen KI ab 2024/25 als redaktionelle Infrastruktur etabliert wurde.⁵⁹ Damit stabilisiert KI kurzfristig die Produktionsfähigkeit, erhöht aber zugleich den publizistisch-personellen Anpassungsdruck auf Redaktionen. Die redaktionelle Souveränität bleibt abhängig von externen Plattformen, proprietären Modellen und konzernstrategischen Entscheidungen – ein Kennzeichen ökonomisch getriebener, aber strukturell fraglicher KI-Resilienz.

www.welt.de

Die drei Mini-Case-Studies machen zentrale Muster instrumentell-ökonomisch getriebener KI-Resilienz sichtbar. Erstens entsteht Resilienz in diesen Fällen hauptsächlich über Effizienzgewinne und Skalierung, nicht über institutionelle Absicherung. Zweitens bleibt sie stark abhängig von Marktbedingungen, Technologiepartnerschaften und unternehmerischen Entscheidungen und ist damit strukturell fragil. Drittens verschiebt sich Verantwortung im Alltag häufig in Richtung technischer Vorstrukturierung, wer KI-Ergebnisse prüft, wer bei Zweifeln entscheidet und wer den Einsatz eines Systems aussetzt, wenn Qualitätsprobleme oder Fehler auftreten.

Ökonomisch getriebene KI-Resilienz kann daher nur kurzfristig Handlungsfähigkeit sichern, verstärkt jedoch mittelfristig den allgemeinen Anpassungsdruck auf Redaktionen. Wo kulturelle Verankerung, Kompetenzaufbau und klare Zuständigkeiten fehlen, wird KI in dieser Hinsicht eher zum Verstärker bestehender Verwundbarkeiten. Die Beispiele zeigen damit präzise die Grenze dieses Resilienztyps: Er stabilisiert Produktion, aber nicht notwendigerweise das professionelle Urteilsvermögen oder die organisationale Lernfähigkeit.

D. Ethik- und vorsorgeorientierte KI-Resilienz:

Selbstbegrenzung, Vertrauen und professionelle Identität

Eine weitere Ausprägung journalistischer KI-Resilienz zeigt sich in Redaktionen, die den technologischen Wandel nicht primär über Tempo oder Skalierung beantworten, sondern über *Prävention*. In diesen Fällen steht nicht die Frage im Vordergrund, was KI leisten kann, sondern unter welchen Bedingungen ihr Einsatz mit journalistischer Verantwortung, Glaubwürdigkeit und professioneller Identität konform geht. Resilienz entsteht hier durch bewusste Entschleunigung, normative Reflexion und die professionelle Entscheidung, bestimmte Formen der Automatisierung nicht oder nur sehr selektiv zuzulassen – Stichwort: Selbstbegrenzung.

- Beim US-amerikanischen Radiosender NPR („National Public Radio“) wird dieses Muster besonders deutlich. Der Umgang mit KI war dort vor einigen Jahren noch von starker Zurückhaltung geprägt. Bevor neue Anwendungen breit eingesetzt werden, wird ihre Wirkung auf Vertrauen, redaktionelle Unabhängigkeit und Publikumsbeziehung geprüft. Diese Vorsicht ist zum einen das Ergebnis eines stark auf Glaubwürdigkeit und journalistische Integrität trainierten Selbstverständnisses. KI wird von den NPR-Verantwortlichen deshalb nicht zwingend als Innovationsmotor gesehen, sondern als potenzielle Irritationsquelle, deren Folgen abgewogen werden müssen, bevor sie (breiter) eingesetzt wird. Konkret nutzt NPR die KI bislang vor allem unterstützend, etwa zur Recherchehilfe, zur Transkription und zur sprachlichen Überarbeitung interner Textentwürfe, nicht jedoch zur autonomen Erstellung redaktioneller Inhalte. Resilienz entsteht hier prinzipiell aus der Fähigkeit, Risiken frühzeitig zu reflektieren und Handlungsoptionen offen zu halten, statt sich vorschnell festzulegen.⁶⁰ Die jüngsten politischen Entwicklungen rund um die Schließung der Corporation for Public Broadcasting (CPB) machen allerdings deutlich, dass NPRs vorsorgender Umgang mit KI auch aus einer strukturell prekären institutionellen Lage heraus erfolgt.⁶¹ Damit zeigt sich: Resilienz kann nicht vor struktureller Verwundbarkeit gegenüber politischen Eingriffen schützen.

www.npr.org

- Bei der „Süddeutschen Zeitung“ zeigt sich eine präventive KI-Resilienz vor allem im praktischen Umgang mit Unsicherheiten bezüglich der KI im Redaktionsalltag: KI wird bisher – wenn überhaupt – nur sehr punktuell z.B. bei der Recherche oder als Strukturierungshilfe genutzt. Ergebnisse aus der Arbeit mit KI-Systemen gelten ausdrücklich als Arbeitsmaterial wie andere Werkzeuge auch. Die Verantwortung für Auswahl, Gewichtung und Veröffentlichung verbleibt grundsätzlich bei den Redakteur:innen. Der KI-Einsatz erfolgt bewusst nicht standardisiert, sondern bleibt an situative Entscheidungen gebunden, die je nach Thema und Form variieren. So stehen laut Selbstauskunft nicht Effizienzgewinne im Vordergrund, sondern Kontrolle und Qualitätssicherung von Menschenhand. Resilienz entsteht hier – wenn man der redaktionellen Leitlinie folgt – durch die klare Haltung, KI nicht als obligatorisches Arbeitsmittel zu verstehen, sondern als Zusatzangebot, das von der Redaktion angenommen oder abgelehnt werden kann. Genau darin liegt der vorsorgende Charakter dieses Praxismusters. Unsicherheiten werden nicht durch automatisierte Prozesse kaschiert. Stattdessen bleiben sie Teil (menschlicher) redaktioneller Entscheidungen und sollen in der Berichterstattung transparent gemacht werden.⁶²

www.sueddeutsche.de

- Beim „Schweizer Radio und Fernsehen“ (SRF) wird KI-Resilienz ebenfalls vor allem über präventive Selbstbegrenzung und normative Klarheit organisiert. Das SRF hat sich im Februar 2024 eigene Regeln zum Umgang mit KI in seine grundlegenden publizistischen Leitlinien geschrieben, die Verantwortung, Transparenz und journalistische Sorgfaltspflicht als zentrale Prinzipien formulieren.⁶³ Demnach dürfen KI-Tools neben der Unterstützung von Recherche oder der Strukturierung von Inhalten auch zur Bearbeitung von Audio- und Videomaterial genutzt werden, allerdings nur, wenn alle Ergebnisse journalistisch verifiziert und in den redaktionellen Entscheidungsprozess eingebettet werden. Bei KI-generierten oder bearbeiteten Medieninhalten verpflichtet das SRF zur öffentlichen Kennzeichnung, damit das Publikum erkennen kann, wo KI-Unterstützung stattgefunden hat. Es hat damit einen neuen instrumentellen Rahmen definiert, in dem die publizistische Verantwortung und Entscheidungshoheit praktisch stets bei den Redaktionen verbleiben. Resilienz soll hier aus dem permanenten Abgleich zwischen realer technologischer Nutzung, redaktioneller Verantwortung und dem Vertrauen gegenüber dem Publikum entstehen.

www.srf.ch

Gemeinsam ist diesen Beispielen, dass KI nicht als Treiber redaktioneller Transformation begriffen wird, sondern als Prüfstein journalistischer Verantwortung. NPR, „Süddeutsche

Zeitung“ und SRF reagieren auf KI gerade nicht mit entfesselter Technik-Euphorie oder Skalierungsfantasien. Vielmehr setzen sie die bewusste Entscheidung, den redaktionellen Einsatz klar zu begrenzen, zu verzögern oder an situative Entscheidungen rückzubinden. Digitale Resilienz entsteht in diesen Fällen durch die Grundsatzfrage, an welchen Stellen KI im Journalismus unterstützend eingesetzt werden kann und wo sie journalistische Prinzipien zu verletzen droht. Die Beispiele zeigen somit, dass Prävention keine passive Haltung sein muss.

Zugleich machen die Fälle deutlich, dass ethik- und vorsorgeorientierte KI-Resilienz kein universell tragfähiges Modell ist. Sie schützt professionelle Identität und publizistisches Vertrauen, bleibt jedoch abhängig von institutionell-wirtschaftlicher Stabilität und organisatorischem Rückhalt. Wo politische, ökonomische oder strukturelle Unsicherheiten zunehmen – wie im Fall NPR – stoßen auch präventive Maßnahmen an ihre Grenzen. Dennoch zeigt sich: In einem Umfeld beschleunigter technologischer Entwicklung stellt gerade diese Form der kontrollierten Öffnung eine eigenständige Resilienzdimension dar.

E. Experimentelle und aufmerksamkeitsgetriebene KI-Resilienz: Sichtbarkeit, Symbolik und offene Folgen

Ein fünftes Praxismuster journalistischer KI-Resilienz zeigt sich dort, wo KI vor allem als Mittel zur Sichtbarkeit von Innovationsleistung eingesetzt wird. In diesen Fällen steht weniger die interne Organisation redaktioneller Arbeit im Vordergrund als die *Inszenierung technologischer Modernität* einzelner Redaktionen. Resilienz wird hier vorrangig über öffentliche Innovationsnarrative angestrebt, nicht über institutionelle Kontrolle oder systematischen Kompetenzaufbau. KI fungiert vielmehr als Signal in die Branche, mit dem Medienhäuser als *Early Adopters* oder *First Movers* zuweilen ihre Anschlussfähigkeit an eine beschleunigte Medienöffentlichkeit coram publico demonstrieren wollen.

- Bei der „Los Angeles Times“ wurde dieses Muster im Jahr 2025 besonders deutlich. Mit dem experimentellen Feature „Insights“ testete das Haus öffentlich sichtbar den Einsatz generativer KI im Meinungsressort. Unter ausgewählten Kolumnen wurden KI-basierte Zusammenfassungen sowie automatisch generierte Gegenperspektiven eingeblendet, die den Anspruch verfolgten, politische Einordnung und Perspektivenvielfalt zu fördern. Das Experiment war bewusst als Test angelegt und für das Publikum unmittelbar erkennbar.⁶⁴ Gleichzeitig machten inhaltliche Fehler und problematische Gewichtungen schnell deutlich, dass *KI-gestützte Perspektivenerzeugung* selbst dort, wo sie nicht als journalistische Autorenschaft auftritt, die Perzeption redaktioneller Verantwortung berührt.⁶⁵ Resilienz entsteht hier zwar kurzfristig durch Innovationsbereitschaft und öffentliche Experimentierfähigkeit, bleibt jedoch unzureichend, weil

redaktionelle Einbettung und die Absicherung von KI-Funktionen nur begrenzt ausformuliert ist.

www.latimes.com

- Eine noch stärker konfliktbehaftete Variante experimenteller KI-Resilienz zeigt sich bei „Politico“. Das Politikmedium setzte 2024 und 2025 KI-gestützte Tools unter anderem zur automatisierten Zusammenfassung politischer Großereignisse und zur Erstellung datenbasierter Analyseformate ein, teils in unmittelbarer Nähe zur laufenden Berichterstattung. Diese Anwendungen waren erkennbar in einzelne Produkte und Dienste integriert und wurden offensiv als Innovationsmodus kommuniziert. Zugleich führten sie zu teils heftigen internen Auseinandersetzungen über redaktionelle Standards, Zuständigkeiten und die Frage, inwieweit experimentelle KI-Produkte denselben publizistischen Maßstäben genügen können wie die ausschließlich von Menschen gemachten.⁶⁶ Resilienz kann in diesem Fall aus der schnellen Erprobung neuer Formate und einem produktiven Einsatz erwachsen. Gleichzeitig wird deutlich, dass diese Form der Resilienz vor allem dann Resonanz erzeugt, wenn redaktionelle Klärungen mit der Geschwindigkeit von KI-Experimenten nicht mithalten.

www.politico.com

- Das Duisburger Lokalfernsehen „Studio 47“ setzt in offensiver Regelmäßigkeit KI in der Produktion und im Sendebetrieb ein, um Effizienz und Reichweite unter strukturellem Druck aufrechtzuerhalten. Als lokaler Fernsehsender mit begrenzten personellen und finanziellen Ressourcen nutzt „Studio 47“ selbst entwickelte KI-Technologie als durchgehendes Stil- und Arbeitsmittel im redaktionellen Alltag.⁶⁷ Vor allem der Einsatz KI-generierter Avatare in der Nachrichtenmoderation ist zentral und wegweisend: Die Avatare übernehmen Moderationsaufgaben und sichern Kontinuität im Programm bei gleichzeitigem Personalmangel.⁶⁸ Der KI-Einsatz wird transparent kommuniziert und ist Teil der öffentlichen Selbstdarstellung des Senders. KI wird zudem als Alleinstellungsmerkmal technischer Innovationsfähigkeit gegenüber Publikum, Förderinstitutionen und potenziellen Kooperationspartner:innen kommuniziert.⁶⁹ Der Case zeigt, dass KI-Experimente konsequent und transparent betrieben werden können, zugleich aber strukturell abhängig bleiben von Förderlogiken, Marktbedingungen und technischer Eigenentwicklung. KI stabilisiert kurz- und mittelfristig den redaktionellen Betrieb, ersetzt jedoch keine langfristige institutionelle Absicherung.

www.studio47.de

- Einen Grenzfall dieses Praxismusters markiert das US-amerikanische Start-up „Channel 1“, das bislang vor allem Demo- und Pilotformate veröffentlicht, in denen generative KI-Avatare als Nachrichtensprecher:innen auftreten und Inhalte aus bereits bestehenden journalistischen Quellen präsentieren. KI wird hier nicht lediglich als eines von vielen Werkzeugen im Hintergrund eingesetzt, sondern wird selbst zum sichtbaren Akteur eines Nachrichtenangebots, das der Sender aktiv vermarktet. „Channel 1“ stellt derzeit keinen journalistischen Betrieb im engeren Sinne dar, sondern ein experimentelles Projekt, das KI-gestützte Nachrichtenproduktion primär technologisch demonstriert.⁷⁰ Das Unternehmen kommuniziert zugleich die Perspektive, ein skalierbares, KI-basiertes Nachrichtenprodukt bis zur Marktreife entwickeln zu können. Darin liegt vor allem eine Signalwirkung für Investor:innen und Zukunftsnarrative. Gerade dadurch markiert „Channel 1“ eine wichtige Grenze experimenteller KI-Resilienz: Solange journalistische Verantwortung, Organisationsstrukturen und publizistische Praxis nicht operationalisiert sind, bleibt völlig offen, ob aus technologischer Versiertheit tatsächlich journalistische Handlungsfähigkeit entsteht.

www.channel1.ai

Gemeinsam ist diesen Beispielen, dass KI-Resilienz hier primär über Experimentiergeist und symbolische Innovationsagitation erzeugt wird. Diese Form der Resilienz ist demnach agil und reaktionsschnell, strukturell aber begrenzt. Sie kann kurzfristig helfen, Aufmerksamkeit zu generieren und technologische Anschlussfähigkeit zu inszenieren. Langfristig bleibt offen, ob solche Experimente ohne *institutionelle Erdung* dazu beitragen, journalistische Autorität und professionelle Verantwortung zu restabilisieren. Die Beispiele markieren damit weniger einen Endpunkt als eine Übergangszone: Sie zeigen, wie und wo sich durch Innovationsbereitschaft ein Momentum entfaltet – und zugleich, wo organisationale Resilienz erst stringent ausgebildet werden muss.

3.5 Resilienz: professionelle Selbstvergewisserung im KI-Zeitalter

Die zuvor beschriebenen Beispiele zeigen, dass journalistische KI-Resilienz nicht automatisch aus der Einführung neuer Prozesse, Regeln oder Technologien entsteht. Resilienz entwickelt sich vielmehr dort, wo Organisationen und Journalist:innen diese Strukturen im Alltag sinnvoll nutzen, hinterfragen und weiterentwickeln. Sie ist damit kein Gegensatz zu Prozessen, sondern zeigt sich darin, wie mit ihnen *konkret* gearbeitet wird. Resilienz ist ein fortlaufender (souveräner) Umgang mit Unsicherheit – nicht ein Zustand, der einmal erreicht und dann gesichert ist.

Diese professionelle Selbstvergewisserung wird durch KI besonders herausgefordert. Denn KI verändert nicht einfach Abläufe, sie verschiebt, wie journalistische Entscheidungen vorbereitet werden: durch automatische Vorschläge, Vorstrukturierungen oder Priorisierungen. Wichtig ist dabei eine begriffliche Klarheit: Nicht jede Automatisierung ist KI, und nicht jede technische Vorstrukturierung greift tief in publizistische Entscheidungen ein. Viele heute genutzte Systeme – etwa zur Transkription, Sortierung oder Themenbeobachtung – unterstützen Arbeitsschritte, ohne selbst journalistische Urteile fällen zu können. Die Herausforderung entsteht aber dort, wo solche Systeme bereits im Vorfeld Einfluss darauf nehmen, was gesehen, gewichtet oder weiterverfolgt wird.

Resilienz bedeutet vor diesem Hintergrund nicht, publizistische Kontrolle an Technologie abzugeben, sondern bewusst damit umzugehen, dass technische Systeme bereits an der Vorbereitung von Entscheidungsprozessen beteiligt sind. Resiliente Redaktionen nutzen daher technische Unterstützung, ohne ihre eigene Urteilskraft aus der Hand zu geben. Sie akzeptieren, dass Entscheidungen nicht mehr bei null beginnen, bestehen aber darauf, dass sie begründet, überprüft und verantwortet werden.

Dabei geht es weniger um eine abstrakte ‚Haltung‘ im Sinne einer inneren Einstellung als um eine praktische Orientierung im Arbeitsalltag. Resilienz zeigt sich darin, dass sich Redaktionen regelmäßig fragen, wie sie mit neuen Werkzeugen arbeiten wollen, wo sie ihnen folgen und wo sie widersprechen. Diese Selbstverständigung ist kein einmaliger Akt, sondern Teil laufender redaktioneller Praxis – etwa, wenn Entscheidungen erklärt, Fehler offen benannt oder technische Annahmen kritisch geprüft werden.

Die Praxisbeispiele zeigen, dass Resilienz dabei keine defensive Verweigerung von Innovation ist, sondern – im Gegenteil –, dass sie dort besonders tragfähig ist, wo Journalismus weder unkritisch auf technologische Versprechen reagiert noch reflexhaft blockiert. Resiliente Redaktionen halten Ambivalenz aus: KI kann entlasten und zugleich neue Risiken erzeugen; sie kann Qualität unterstützen und diese gleichzeitig unter Druck setzen.

In diesem Sinne bedeutet KI-resilient zu sein auch, Unsicherheiten als Element journalistischer Arbeit im digitalen Raum bis zu einem gewissen Maß auszuhalten. Redaktionen und Medienschaffende, die resilient sind, reagieren darauf mit Reflexion statt mit Desorientierung. Damit unterscheidet sich Resilienz grundlegend von rein ökonomischen oder technologischen Überlebensstrategien. Während diese vor allem auf kurzfristige Stabilisierung zielen, richtet sich Resilienz auf die langfristige Sicherung journalistischer Integrität. Sie fragt, wie sich der Journalismus unter KI-Bedingungen weiterentwickeln kann, und wie er trotzdem seinem eigenen Anspruch gerecht bleibt.

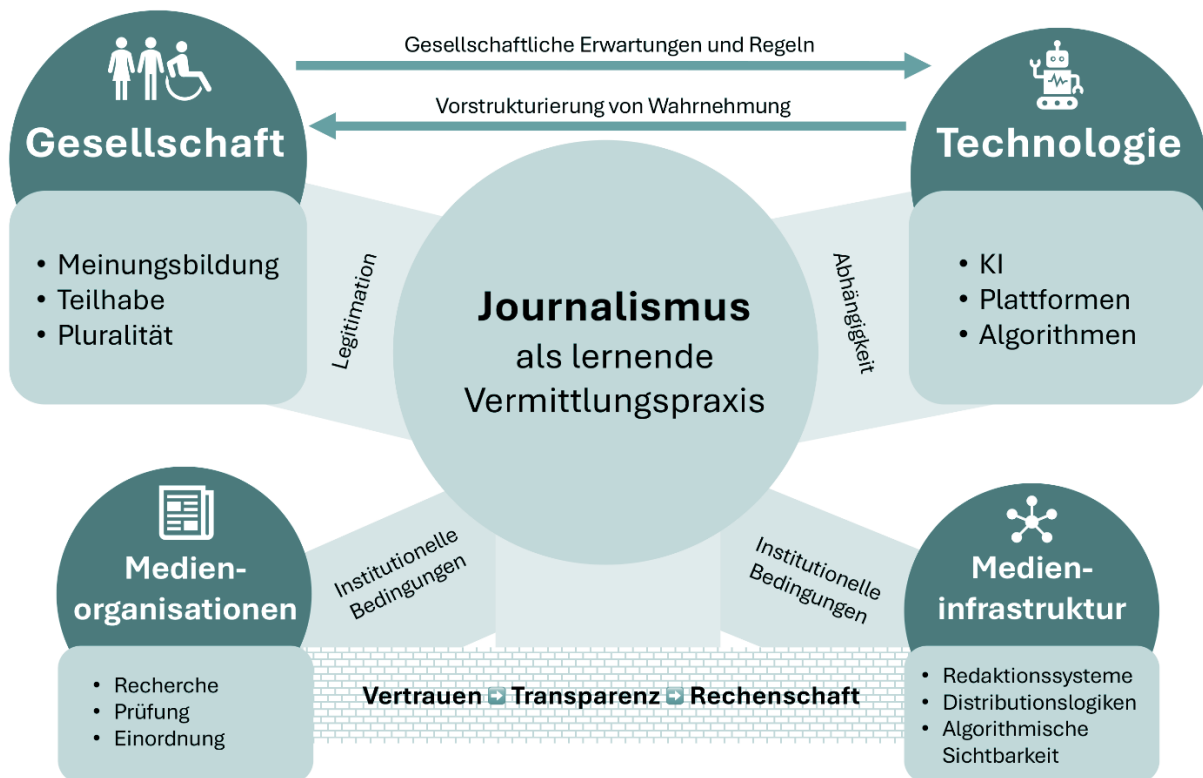


Abbildung 8: Journalismus als vermittelnde Praxis digitaler Öffentlichkeit (eigene Darstellung)

Die bisherigen Kapitel haben gezeigt, dass KI den Journalismus nicht nur technisch, sondern strukturell herausfordert. Abbildung 8 zeigt, wie Journalismus als Kreislauf zwischen ‚Gesellschaftsleistung‘ (Meinungsbildung) und ‚Technologie/Plattformen‘ (Algorithmen) fungiert, wobei Vertrauen und Transparenz zentrale Vermittlungsbedingungen sind. Die hierfür relevanten Fragen nach Verantwortung, Urteilkraft und redaktioneller Souveränität lassen sich nicht allein durch individuelle Kompetenz oder organisationale Praxis beantworten. Sie verlangen nach kollektiven Setzungen, die orientieren, Grenzen markieren und gemeinsame Erwartungen formulieren. Genau an diesem Punkt setzen die im Folgenden diskutierten Leitlinien, Regelwerke und Positionspapiere an.

In den vergangenen Jahren – beschleunigt durch den weltweiten Durchbruch und die großen Schritte bei der Weiterentwicklung generativer KI – haben zahlreiche Redaktionen, Medienhäuser, Berufsverbände und internationale Initiativen normative Rahmenwerke für den Einsatz von KI im Journalismus vorgelegt. Diese Dokumente reagieren auf eine verbreitete Erfahrung von Kontrollverlust und Verunsicherung. Sie sollen klären, was im Umgang mit KI gilt, welche Praktiken als gewinnbringend oder zielführend angesehen werden und welche nicht. Solche Leitlinien fungieren damit als Varianten institutioneller

Selbstverpflichtung für Situationen, in denen technologische Entwicklung schneller voranschreitet als die etablierten Routinen journalistischer Standardisierung.

Im Folgenden nehmen wir diese Leitlinien nicht als finale Lösungen in den Blick, sondern als Ausdruck einer zum Teil dynamischen Suchbewegung, einem fortlaufenden Erneuerungsprozess in einer durch generative KI-Systeme rasant beschleunigten Medienwelt. Wir fragen, welche Probleme die Papiere adressieren, welche Vorstellungen von Verantwortungszuschreibung und Urheberschaft ihnen zugrunde liegen – und welche Aspekte journalistischer Resilienz sie bislang ein- und ausblenden. Dabei geht es ausdrücklich nicht um eine Abwertung einzelner Dokumente, sondern um eine systematische Einordnung ihrer Fokusse, Gemeinsamkeiten und möglichen Leerstellen.

Nach unserem Grundverständnis verstehen wir die kollektive und kooperative Arbeit an Leitlinien im Medienbetrieb als spezifische Resilienzstrategie – nämlich als Versuch, Stabilität durch normative Orientierung herzustellen. Zugleich zeigt sich, dass Leitlinien vor allem das *Welche Regeln gelten?* beantworten sollen, während Fragen des *Wie wird im Alltag (unter Zeitdruck) entschieden?* häufig offenbleiben. Diese Spannung ist zentral für das Verständnis ihrer Wirkungen – aber auch ihrer Grenzen.

4. Leitlinien, Regelwerke und Positionspapiere: Orientierung, Selbstbindung und ihre Grenzen

Worauf dieses Kapitel Antworten gibt:

1. Warum Leitlinien im Journalismus als Reaktion auf Unsicherheit, Beschleunigung und Verantwortungsdiffusion entstehen.
2. Welche normativen Muster internationalen und deutschen KI-Leitlinien im Journalismus zugrunde liegen.
3. Wo Leitlinien Orientierung geben – und wo sie an Grenzen stoßen, weil sie Alltagspraxis nicht vorausschauend strukturieren.
4. Welche übertragbaren Logiken sich aus anderen regulierten Medienbereichen ableiten lassen.
5. Warum bisherige Leitlinien notwendige, aber nicht hinreichende Instrumente resilienter journalistischer Praxis sind.

4.1 Warum Leitlinien?

Normative Antworten auf strukturelle Unsicherheit

Leitlinien sind Reaktionen auf Situationen, in denen bestehende Routinen, implizite Standards und professionelle Selbstverständlichkeiten meist nicht ausreichen, um Handlungsfähigkeit und Legitimität zu sichern. Im Kontext des journalistischen KI-Einsatzes sollen sie helfen, mit Unsicherheiten umzugehen und für den Umgang mit Entscheidungserwartungen zu sensibilisieren – nach innen wie nach außen. Derzeit werden Entscheidungen zunehmend technisch vorbereitet,⁷¹ Prozesse intransparent vorstrukturiert,⁷² und Verantwortlichkeiten über Mensch-Maschine-Konstellationen hinweg verteilt.⁷³ Leitlinien sollen hier eine mögliche Orientierung bieten, wo situatives Entscheiden alleine nicht ausreicht.

In analytischer Perspektive lassen sich Leitlinien als *normative Selbstbindung* verstehen: Redaktionen und Medienhäuser formulieren auf dieser Basis Erwartungen an sich selbst und ihre Mitarbeitenden. Sie definieren Grenzen des Zulässigen, benennen Verantwortlichkeiten und markieren rote Linien, etwa im Hinblick auf die Urheberschaft, Transparenz oder Art und Umfang des Einsatzes generativer KI. Solche Leitlinien sind damit weniger operative Handlungsanweisungen als vielmehr *kodifizierte Selbstbeschreibungen* im

Sinne von berufsethischen Verhaltensgrundsätzen;⁷⁴ denn sie schreiben fest, wofür eine Organisation stehen will, wenn technologische Dynamiken deren etablierte Ordnungsvorstellungen aus den Angeln zu heben drohen.

Dieses Verständnis von Leitlinien erklärt, warum sie im Journalismus häufig stark prinzipienorientiert und meist sehr abstrakt formuliert sind. Sie artikulieren Werte, betonen professionelle Normen und knüpfen an bestehende Ethikdiskurse an – beispielhaft stehen dafür der gemeinsame KI-Kodex der ARD-Landesrundfunkanstalten, Deutsche Welle, ZDF und Deutschlandradio sowie noch einmal gesondert die „KI-Grundsätze des ZDF“ (vgl. Kap. 4.3). Gerade im Umgang mit KI dienen sie dazu, Kontinuität sicherzustellen: Journalistische Standards und Werte sollen demnach auch unter veränderten technischen Bedingungen erhalten bleiben.

Zugleich tun sich hier erste Gräben auf: Leitlinien beantworten zumeist die Frage, welche Regeln gelten sollen, nicht aber, wie diese im redaktionellen Alltag konkret umzusetzen sind. Sie bleiben einerseits oft bewusst nebulös, um für unterschiedliche redaktionelle Kontexte anschlussfähig zu sein, entziehen sich dadurch jedoch genau jenen Situationen, in denen Resilienz praktisch angezeigt wäre: beispielsweise unter Zeitdruck, bei unklaren Zuständigkeiten oder im Umgang mit Fehlern, die nicht eindeutig zuordenbar sind.⁷⁵

Andererseits sind die bisherigen Regelwerke immer unvollendet und unterliegen einem dynamischen Anpassungsprozess: Vor diesem Hintergrund sind Leitlinien angesichts der rasanten technologischen Fortschritte im KI-Bereich als temporäre Maßnahme zu verstehen und markieren den Versuch, Verantwortung und Selbstverständnis in einer Momentaufnahme abzubilden. Ihre Stärke liegt in der groben Orientierung, ihre offenkundige Schwäche in der fehlenden Übersetzungsleistung redaktioneller Praktiken.

4.2 Internationale KI-Leitlinien und Orientierungsrahmen im Journalismus: Muster, Schwerpunkte, Leerstellen

Ein Vergleich internationaler Leitlinien und Orientierungsrahmen zum Einsatz von KI im Journalismus zeigt trotz unterschiedlicher institutioneller Herkunft eine bemerkenswerte inhaltliche Nähe. Analysiert wurden sowohl ausgewählte redaktionelle und unternehmensbezogene Regelwerke einzelner Medienhäuser und Nachrichtenagenturen als auch branchenweite Selbstverpflichtungen und übergeordnete normative Rahmen aus dem internationalen Kontext. Festzustellen ist, dass trotz dieser Vielfalt zentrale Prinzipien in nahezu identischer Form wiederkehren. Diese – positive – Konvergenz verweist auf einen geteilten Problemdruck und ein gemeinsames Verständnis dessen, was im KI-Zeitalter als schützens- oder zumindest als hervorhebenswert gilt.

Die Analyse stützt sich auf eine vergleichende Auswertung einer Auswahl internationaler Redaktionsleitlinien zum Einsatz von KI im Journalismus, darunter redaktionelle Regelwerke und Positionspapiere von „The Associated Press“ (USA), APA (Österreich), BBC sowie „The Guardian“ (Großbritannien), Mediahuis (Belgien/Niederlande), NZZ, Ringier und SRF (Schweiz) sowie der New York Times Company (USA).

1. Zu den gemeinsamen Kernthemen zählt an erster Stelle das **Human-in-the-Loop-Prinzip**. In nahezu allen untersuchten redaktionellen Leitlinien wird ausdrücklich festgehalten, dass journalistische Verantwortung nicht automatisiert werden darf und redaktionelle Entscheidungen beim Menschen verbleiben müssen. Diese Setzung findet sich sowohl in agenturbezogenen Regelwerken wie jenen der „Associated Press“ oder der APA als auch in redaktionellen Leitlinien einzelner Medienhäuser, etwa bei „New York Times“, „Guardian“, der BBC, SRF, NZZ, Ringier oder dem Mediahuis-Konzern. KI wird hier ausdrücklich als unterstützende Infrastruktur begriffen, die an klare Governance- und Verantwortlichkeitsstrukturen gebunden ist. Damit soll KI weder redaktionelle Urheberschaft noch eigenständige Entscheidungsbefugnisse überlassen werden. Diese Festlegung dient in allen Fällen der Stabilisierung professioneller Identität.
2. Eng damit verbunden ist das zweite Kernthema: **Transparenz und Kennzeichnung**. Nahezu alle untersuchten Leitlinien betonen, dass der Einsatz von KI-Systemen für Publikum und interne Akteure nachvollziehbar gemacht werden muss. Dies reicht von allgemeinen Transparenzforderungen, wie sie etwa beim „Guardian“, bei der „Associated Press“ oder der „New York Times“ formuliert werden, bis hin zu expliziteren Kennzeichnungspflichten, wie sie etwa SRF oder der Schweizer Presserat vorsehen. Auffällig ist jedoch, dass Transparenz trotz ihrer zentralen Bedeutung meist als normatives Ziel formuliert wird. Konkrete Aussagen darüber, wann Kennzeichnungen zwingend erforderlich sind, wie detailliert sie ausfallen sollen oder wie mit Grenzfällen umzugehen ist, bleiben in vielen Leitlinien offen.

In der Zusammenschau lässt sich diese Übereinstimmung als dominante Lesart interpretieren: Die hier ausgewählten redaktionellen Leitlinien zum Einsatz von KI sind primär ethik- und governance-orientiert. Sie zielen darauf ab, Verantwortung festzuschreiben, Risiken zu begrenzen und journalistische Grundwerte zu schützen. Resilienz erscheint vor allem als Abwehr unerwünschter Kollateralschäden und zur Sicherung bestehender professioneller Normen und Standards. Lern-, Anpassungs- oder experimentelle Dimensionen in der journalistischen Praxis mittels und durch KI scheinen hingegen nur ansatzweise auf.

Gerade an dieser Stelle werden typische Limitierungen sichtbar. Leitlinien operieren weiterhin auf der Ebene abstrakter Prinzipien und verzichten auf einen ausgearbeiteten Anwendungsbezug. Exemplarisch zeigt sich diese Begrenzung in den KI-Leitlinien der BBC: Zwar wird dort mehrfach gefordert, dass jeder Einsatz von KI unter aktiver menschlicher redaktioneller Aufsicht stehen müsse und für das Publikum transparent zu klären sei. Was unter angemessener Aufsicht konkret zu verstehen ist, bleibt jedoch offen: Die Leitlinien differenzieren zwar zwischen verschiedenen Einsatzfeldern – etwa zwischen Empfehlungssystemen, datenanalytischer Unterstützung oder der Bearbeitung von Audio- und Bildmaterial –, übersetzen diese Unterscheidung jedoch nicht in verbindliche Zuständigkeiten, Prüfprozesse oder Entscheidungsabläufe. Auch bei der Transparenz bleibt der Bezug abstrakt. Die BBC verlangt, das Publikum kontextangemessen („in a manner appropriate to the context“)⁷⁶ über KI-Nutzung zu informieren, lässt jedoch offen, wie und ab welchem Grad der KI-Unterstützung eine solche Kennzeichnung verpflichtend ist, in welcher Form sie zu erfolgen hat und wer im Zweifel darüber entscheidet.

Redaktionelle Verantwortung wird damit zwar eindeutig beim Menschen verortet, ihre praktische Wahrnehmung aber weitgehend an situative Einzelfallentscheidungen delegiert. Die Leitlinien benennen damit zentrale Prinzipien, ohne sie in belastbare Routinen für den redaktionellen Alltag zu übersetzen – also weitestgehend ohne Aussagen zu Alltagssituationen in Redaktionen, etwa bei Zeitdruck, Ressourcenkonflikten oder Eskalationsfällen.

Die Paris Charter on AI and Journalism von 2023⁷⁷ nimmt in diesem Gefüge eine besondere, nicht unumstrittene Rolle ein: Sie formuliert einen normativ stark aufgeladenen Anspruch an journalistische Verantwortung im KI-Zeitalter und positioniert diese im Kontext demokratischer Öffentlichkeit. Zugleich haben exponierte Branchenakteur:innen wie die WAN-IFRA die Charta zwar inhaltlich begleitet, sich jedoch bewusst gegen eine formale Unterstützung entschieden.⁷⁸ Damit markiert die Charta weniger einen operativen Standard als einen normativen Referenzpunkt, an dem sich Zustimmung, Abgrenzung und Konfliktlinien innerhalb der Branche ablesen lassen. Als öffentlichkeitswirksames Dokument liegt ihre Stärke in der klaren Positionierung und der globalen Anschlussfähigkeit. Die Charta soll aber zugleich keine konkreten redaktionellen Abläufe regeln. Damit eignet sie sich vor allem als ethischer Kompass denn als ein konkretes Instrument zur Organisation redaktioneller Praxis.

4.3 Deutsche Leitlinien und Positionspapiere: Selbstverständnis, Vorsicht, institutionelle Kultur

Ein Blick auf deutsche Leitlinien und Positionspapiere zum Einsatz von KI im Journalismus zeigt hingegen eine deutlich andere Akzentuierung. Während internationale Leitlinien häufig von Governance-Logiken und institutioneller Steuerung durchdrungen sind, rücken deutsche Regelwerke stärker *journalistisches Selbstverständnis und Berufsrolle* ins Zentrum. Die folgende Analyse stützt sich auf eine vergleichende Auswertung von Leitlinien und Positionspapieren zum Einsatz von KI im Journalismus in Deutschland. Berücksichtigt wurden dabei Dokumente von Netzwerk Recherche und Deutschen Journalisten-Verband (DJV), Kodizes öffentlich-rechtlicher Anbieter wie ARD, ZDF, MDR, Bayerischer Rundfunk, Deutschlandradio und Deutscher Welle, redaktionelle Selbstverpflichtungen überregionaler Medien wie *Der Spiegel*, *Die Welt*, *taz* und dpa sowie Leitlinien und Selbstbeschreibungen zahlreicher Lokal- und Regionalmedien, darunter „Allgemeine Zeitung“, „Kölner Express“, Medienhaus Aachen, „Mindener Tageblatt“, „Rheinpfalz“, der Schleswig-Holsteinische Zeitungsverlag (shz) und „Südkurier“.

Über diese sehr unterschiedlichen Akteur:innen hinweg zeigt sich eine auffällige Übereinstimmung in der Grundhaltung: KI darf journalistische Arbeit unterstützen, aber weder redaktionelle Entscheidungen treffen noch Urheberschaft übernehmen. Diese Abgrenzung wird in den Leitlinien immer wieder an konkreten Tätigkeiten festgemacht: KI kann Interviews transkribieren, Texte zusammenfassen, Übersetzungen vorschlagen oder große Dokumentenbestände durchsuchbar machen. Sie soll jedoch keine Artikel eigenständig verfassen, keine Schlagzeilen festlegen und keine publizistischen Bewertungen vornehmen. Verantwortung, so der wiederkehrende Tenor, bleibt unteilbar menschlich.

Besonders pointiert formulieren Netzwerk Recherche und DJV diese Grenzziehung. In ihren Positionspapieren von 2025 bzw. 2023 steht weniger die Frage im Vordergrund, wie KI produktiv genutzt werden kann, sondern wie sich Urheberschaft, redaktionelle Kontrolle und faire Arbeitsbedingungen gegen eine schleichende Automatisierung absichern lassen.⁷⁹ Warnungen vor automatisierter Texterstellung, intransparenten Trainingsdaten oder der Entwertung journalistischer Arbeit prägen den Ton – dies sicherlich auch vor dem Hintergrund eines vermuteten Stellenabbaus in größerem Maßstab, gegen den der DJV als berufsgenossenschaftliche Interessensvertretung seit einigen Jahren arbeitsrechtlich vorgeht. Lernen erscheint hier primär als individuelle Anforderung an Journalist:innen, etwa mit neuen Recherchewerkzeugen kompetent umzugehen, aber nicht den KI-Einsatz organisationsweit zu gestalten.

Diese professionsethische Rahmung setzt sich in den Leitlinien etablierter deutscher Redaktionen fort, allerdings mit jeweils unterschiedlichem institutionellem Akzent. Die

öffentlich-rechtlichen Sendeanstalten ARD, ZDF, „Deutschlandradio“ und „Deutsche Welle“ binden den Einsatz von KI explizit an ihren publizistischen Auftrag. Der 2026 verabschiedete gemeinsame recht kompakte „KI-Kodex“⁸⁰ hebt sich von früheren, eher disparaten Leitlinien vor allem dadurch ab, dass er nicht nur normative Grenzziehungen wiederholt, sondern den KI-Einsatz explizit an den öffentlich-rechtlichen Versorgungsauftrag, dessen strategische Zielsetzungen und seine organisationale Verantwortung gegenüber dem Publikum insgesamt koppelt. KI wird hier auf weniger als einer Doppelseite ausschließlich dann als zulässig beschrieben, wenn sie einen *nachweisbaren Mehrwert für die Erfüllung dieses Auftrags* erzeugt. Damit wird eine Zweckbindung formuliert, die über abstrakte Ethikbekenntnisse hinausgeht und den KI-Einsatz programmatisch begrenzt.

Auffällig ist dabei zudem die klare Verbindung zwischen Personalisierung und Vielfalt. Der Kodex beschreibt so genannte KI-gestützte Personalisierung ausdrücklich als Möglichkeit, Inklusion zu fördern und einen möglichst barrierefreien Zugang zu öffentlich-rechtlichen Inhalten zu eröffnen. Gemeint sind dabei vor allem niedrigschwellige Zugangsformen wie automatisierte Untertitelung, Audiodeskription, Übersetzungen oder adaptive Darstellungsformate, die Inhalte an unterschiedliche sprachliche, kognitive oder körperliche Voraussetzungen anpassen können.

Gleichzeitig benennt der Kodex die Gefahr diskriminierender Trainingsdaten ausdrücklich und versteht deren Bearbeitung nicht als bloße Risikoabwägung, sondern als aktive Gestaltungsaufgabe. Die beteiligten Anstalten verpflichten sich damit implizit zu einem bewussten Umgang mit Verzerrungen in Daten und Modellen, etwa durch Prüfung von Trainingsgrundlagen, kritische Bewertung von Outputs und korrigierende Eingriffe in algorithmische Systeme.⁸¹ KI-gestützte Personalisierung soll so nicht zur Verengung von Perspektiven führen, sondern gezielt eingesetzt werden, um Vielfalt sichtbar zu machen und Filterblasen aktiv entgegenzuwirken. Damit adressiert der KI-Kodex ein Spannungsfeld, das viele journalistische Leitlinien nur randständig behandeln: den gleichzeitigen Anspruch, algorithmische Systeme produktiv zu nutzen und ihre strukturellen Verzerrungen nicht nur zu begrenzen, sondern gestaltend zu bearbeiten.

Insgesamt markiert der KI-Kodex einen wichtigen Entwicklungsschritt. Er bleibt fest in einer öffentlich-rechtlichen Auftragstradition verankert, erweitert diese jedoch um konkrete strategische, kompetenzbezogene und kooperative Elemente. Gleichwohl zeigt sich auch hier die charakteristische Grenze von redaktionellen Leitlinien in Deutschland: Trotz erhöhter Konkretion auf der Ebene von Prinzipien und Zielsetzungen bleibt offen, wie diese Vorgaben in redaktionelle Entscheidungsroutrinen und Alltagspraktiken übersetzt werden können – zumindest sind diese Mechanismen nicht öffentlich zugänglich.

Eine differenziertere Betrachtung ergibt sich am (regionalen) Beispiel des BR. Während frühere Fassungen der BR-Leitlinien den Einsatz von KI vor allem ethisch rahmten und auf allgemeine Prinzipien wie Transparenz, Verantwortung und redaktionelle Kontrolle fokussierten, zeigt die Überarbeitung aus dem Jahr 2024 eine deutliche Verschiebung hin zu prozessualer proaktiver Steuerung – und damit hin zu einem nicht abgeschlossenen KI-Resilienz-Framework (Abbildung 9):⁸² KI wird nicht normativ eingehegt, sondern als potenziell risikobehaftetes Gesamtsystem begriffen, dessen Einsatz organisatorisch auf unterschiedlichen Ebenen bis hin zum Management abgesichert werden sollte. Abnahmeprinzipien, dokumentierte Prüfprozesse, Folgenabschätzungen und experimentelle Schutzräume markieren hierbei einen Übergang von ethischer Selbstvergewisserung zu institutionalisierter Governance. Der BR versteht damit seine Leitlinien als lernende Instrumente und übersetzt diese schrittweise in belastbare Entscheidungsstrukturen.

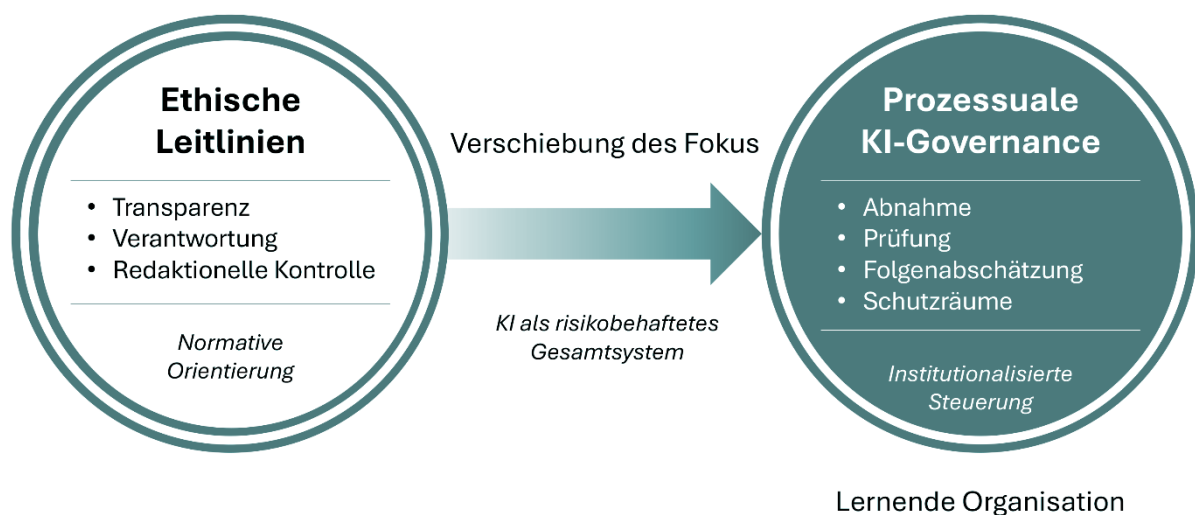


Abbildung 9: Weiterentwicklung der KI-Richtlinien des Bayerischen Rundfunks von 2020 bis 2024: Fokusverschiebung von primär ethischer Selbstbindung zu institutionell verankerter Governance-Logik (eigenen Darstellung)

Bei den privaten überregionalen Medien treten deutlicher abgestufte Positionen zutage. „Der Spiegel“ betont in seiner – nach einer „KI-Panne“ innerhalb der Online-Redaktion rasch erneuerten⁸³ – Selbstbeschreibung vom Herbst 2025 ausdrücklich, dass KI keine Texte „maßgeblich“⁸⁴ verfassen und keine redaktionellen Entscheidungen treffen darf. Gleichzeitig räumt das Haus ein, dass sich deren Einsatzmöglichkeiten rasch verändern. Die eigenen Regeln gelten daher ausdrücklich als vorläufig. In der Praxis bedeutet das, dass KI etwa bei Recherche, Faktenchecks oder internen Analysen zum Einsatz kommen

darf, während Autorenschaft und publizistische Verantwortung beim Menschen verbleiben sollen. Organisationales Lernen ist in dieser Perspektive zwar vorgesehen, aber als kontrollierter Prozess innerhalb fester professioneller Leitplanken.

Die „Süddeutsche Zeitung“ beschreibt ihren Umgang mit KI noch anwendungsnäher. Sie versteht KI als Teil der redaktionellen Infrastruktur und benennt proaktiv konkrete Einsatzfelder: Transkription von Interviews, Übersetzung internationaler Quellen, Moderation von Kommentaren, Unterstützung barrierefreier Angebote. Zugleich hält die SZ konsequent daran fest, dass Entscheidungen über Veröffentlichung, Gewichtung und Tonfall nicht per KI automatisiert werden dürfen. KI soll redaktionelle Arbeit effizienter machen, aber nicht ersetzen. Die Leitlinien verbinden damit operative Offenheit mit klarer Verantwortungszuweisung.⁸⁵

Die „taz“ als eine wichtige Vorreiterin beim Einsatz von KI-generierten Texten – die Zeitung ‚beschäftigte‘ von Herbst 2002 bis zum Frühjahr 2024 mit Anic T. Wae eine:n frühen KI-Kolumnist:in –⁸⁶ wiederum koppelt ihre skeptische Grundhaltung an vergleichsweise präzise Regeln, die im Juli 2025 aktualisiert wurden: Die Redaktion erlaubt KI zur Unterstützung von Recherche, Textüberarbeitung oder Bildbearbeitung, schließt KI-Autorschaft jedoch explizit aus. Auffällig ist, dass die *taz* technische Fragen stärker reflektiert als viele andere Häuser, indem sie etwa den Umgang mit Trainingsdaten, die Wahl von Tools und die Bedeutung von Nachhaltigkeit und Open-Source-Lösungen thematisiert. Auch hier bleibt der Einsatz von KI an journalistische Grundsätze gebunden, zugleich werden die eigenen Regeln ausdrücklich als überprüfbar und veränderbar markiert.⁸⁷

Eine besondere Rolle nimmt die dpa ein. Während frühe Positionspapiere vor allem normative Leitplanken formulierten,⁸⁸ hat die Nachrichtenagentur ihre KI-Strategie seit einigen Jahren deutlich in Richtung praktischer Umsetzung verschoben – wobei deren eigentliche KI-Leitlinien noch immer aus 2023 stammen. Mit dem Programm „Wegweiser KI“ schulte sie gemeinsam mit dem DFKI hunderte Journalist:innen, etwa im Umgang mit Chatbots im Redaktionsalltag oder bei der Erkennung KI-generierter Inhalte. Die daraus entstandene KI-Handreichung aus 2025⁸⁹ übersetzt diese Erfahrungen in konkrete Anwendungsbeispiele, von strukturierter Prompt-Gestaltung bis zur KI-gestützten Recherche. Mit einem KI-Rechercheassistenten setzt die dpa ihre Leitlinien schließlich auch technisch um: Journalist:innen können in natürlicher Sprache im eigenen Archiv recherchieren und erhalten KI-generierte Zusammenfassungen mit direktem Quellenbezug.

Im Lokal- und Regionaljournalismus verdichtet sich diese Logik abermals. Die Leitlinien von Häusern wie „Rheinpfalz“, „Südkurier“, shz, Medienhaus Aachen und „Mindener Tageblatt“ bleiben meist knapp und verweisen vor allem auf journalistische Grundsätze. Sie erlauben KI etwa zur Unterstützung von Recherche oder Textvorarbeit, verzichten jedoch

weitgehend auf Aussagen zu Workflows, Zuständigkeiten oder Prüfprozessen. Der Umgang mit KI bleibt damit stark situationsabhängig und liegt in der Verantwortung einzelner Redakteur:innen.

In der Gesamtschau zeigen deutsche KI-Leitlinien eine hohe normative Klarheit und einen vorsichtigen Zugriff auf KI. Sie benennen präzise, was KI nicht leisten darf und was nicht - und sichern journalistische Verantwortung hierdurch konsequent ab. Zugleich entwickeln einzelne Akteure – insbesondere die dpa und in Teilen auch die öffentlichen-rechtlichen Anstalten (u.a. BR) oder größere Medienhäuser wie der Süddeutsche Zeitung Verlag – konkrete Formen organisationalen Lernens und praktischer Integration. Insgesamt bleibt der Umgang mit KI jedoch relativ eindeutig auf Prävention als auf dynamische Weiterentwicklung ausgerichtet. Die Leitlinien formulieren überzeugend, wofür Journalismus stehen soll; wie Redaktionen unter KI-Bedingungen dauerhaft wandlungsfähig, lernbereit und souverän handeln können, bleibt vielfach offen.

4.4 Zwischen normativer Orientierung und erlebter Praxis: Ergebnisse einer quantitativen Befragung von Journalist:innen

Wie dargelegt beschreiben Leitlinien und Positionspapiere, welche Maßstäbe journalistische Organisationen für den Einsatz von KI in Redaktionen anlegen (würden). Ob und wie diese Marker im Arbeitsalltag tatsächlich praktisch orientieren, lässt sich aus den Leitlinien selbst nur begrenzt erschließen. Gerade bei einer digitalen Innovation, die tief in bestehende Routinen, Rollenselbstbilder und Zukunftsentwürfe des Berufs eingreift, bleibt opak, wie diese Ansprüche von denjenigen wahrgenommen werden, die täglich unter den Bedingungen digitaler Beschleunigung arbeiten.

Vor diesem Hintergrund rückt die Perspektive der Journalist:innen selbst in den Fokus. Die im Folgenden ausgewertete quantitative Befragung von Mitgliedern des DJV aus dem Frühjahr 2025 macht sichtbar, wie KI im redaktionellen Alltag erlebt wird: als Entlastung oder zusätzliche Belastung, als Zukunftschance oder als Grund für Verunsicherung.⁹⁰ Sie zeigt, in welchem Verhältnis zu diesem Zeitpunkt formale Orientierung und praktische Erfahrung stehen und wo sich Spannungen zwischen Anspruch und gelebter Praxis manifestieren. Damit ergänzt die Befragung die Analyse der Leitlinien um eine empirische Innenansicht. Sie liefert Hinweise darauf, welche Voraussetzungen ein resilienter Umgang mit KI jenseits normativer Selbstvergewisserung benötigt.

Die Befragung zeigt deutlich, dass KI im journalistischen Feld vor allem als Teil eines *umfassenderen Transformationsprozesses* wahrgenommen wird. Sie berührt Arbeitsbedingungen, Zukunftserwartungen und das professionelle Selbstverständnis gleichermaßen.

Bereits die Einschätzung zur langfristigen Bedeutung des Journalismus verweist auf eine spürbare Verunsicherung. Auf die Frage, welche Rolle Journalismus in zehn Jahren im Medienrepertoire der Mediennutzenden spielen wird, erwarten 43,6 Prozent der Befragten eine eher untergeordnete Bedeutung. 34,7 Prozent gehen von einer gleichbleibenden Rolle aus, während nur 21,7 Prozent mit einer größeren Bedeutung rechnen (n=1.082). KI erscheint in diesem Zusammenhang als ein weiterer Faktor, der unter Journalist:innen einen schon länger empfundenen schleichenden Bedeutungsverlust ihres Berufsfeldes verstärkt.

Diese Einschätzung korrespondiert mit den Sorgen im Hinblick auf KI in der öffentlichen Kommunikation. 63,9 Prozent der befragten Journalist:innen geben an, sich angesichts der fortschreitenden Verbreitung von KI Sorgen zu machen; 26,9 Prozent äußern keine Sorgen, 9,2 Prozent sind unentschieden (n=934). Damit prägt eine skeptische Grundhaltung die Wahrnehmung einer deutlichen Mehrheit. Zugleich ist diese Haltung differenziert. Das zeigt sich bei der Frage nach persönlicher Verunsicherung durch den Einsatz von KI im Journalismus: Auf einer fünfstufigen Skala stimmen 33,2 Prozent der Aussage „eher“ oder „voll zu“, dass KI sie verunsichert (Wert 4/5). 40,7 Prozent widersprechen „eher“ oder „deutlich“ (Wert 1/2), 24,7 Prozent ordnen sich im Mittelfeld ein; 1,4 Prozent wählen „weiß nicht“ (n=938). KI wird somit weder durchgängig als Bedrohung noch als unproblematische Entwicklung wahrgenommen, sondern als Quelle anhaltender Unsicherheit.

Welche Aspekte dabei besonders beunruhigen, wird in der nachfolgenden Rangfrage sichtbar (Abb. 10). An erster Stelle nennen die Befragten (n=587) den Zuwachs an Desinformation, gefolgt vom Vertrauensverlust in den Journalismus. Ebenfalls hoch, aber nachrangiger gewichtet werden Auswirkungen auf das journalistische Handwerk, ethische Bedenken sowie die Sorge vor Arbeitsplatzverlust. Diese Rangfolge verdeutlicht, dass sich die Sorgen weniger auf individuelle Rationalisierungseffekte als auf strukturelle Folgen für Qualität, Glaubwürdigkeit und professionelle Standards richten.

Was bereitet Ihnen die größten Sorgen?

Anzahl Antworten: 587

Rang	Werte	Rangverteilung	Punkte	Wertungen
1.	Zuwachs an Desinformation		2125	575
2.	Vertrauensverlust in Journalismus		1727	551
3.	Auswirkungen auf das journalistische Handwerk		1330	507
4.	Ethische Bedenken		1146	482
5.	Sorge vor Arbeitsplatzverlust		939	466
		Niedrigste  Höchste		

Abbildung 10: Sorgen in Bezug auf die fortschreitende Verbreitung von Künstlicher Intelligenz in der öffentlichen Kommunikation (eigene Darstellung)

Besonders aussagekräftig für die Frage journalistischer Resilienz sind die Erwartungen an zukünftige berufliche Anforderungen im Journalismus unter den Bedingungen einer fortschreitenden Verbreitung von KI-Technologie in der öffentlichen Kommunikation: In der Mehrfachnennungsfrage zu künftig stärker geforderten Rollen im Journalismus (n=958) wird mit großem Abstand das Fact Checking genannt: 87,8 Prozent sehen hier einen wachsenden Bedarf. Es folgen die Vermittlung von Nachrichten- und Medienkompetenz (68,0 Prozent), das Aufzeigen von Perspektiven und Lösungsszenarien (60,0 Prozent) sowie kuratierende Tätigkeiten (54,0 Prozent). Moderation (35,5 Prozent) und explizites Engagement für gesellschaftlichen Zusammenhalt (37,2 Prozent) werden im direkten Vergleich seltener genannt (Abb. 11). Aus Sicht der Befragten verschiebt sich die berufliche Anforderungslogik damit in Richtung Prüfung, Einordnung und erklärender Vermittlung. Der professionelle Mehrwert wird vor allem dort gesehen, wo *menschliches Urteilsvermögen und Verantwortung* gefragt bleiben.

In welchen Bereichen bzw. in welchen Rollen werden Journalist:innen künftig mehr gefordert sein?

Anzahl Antworten: 958

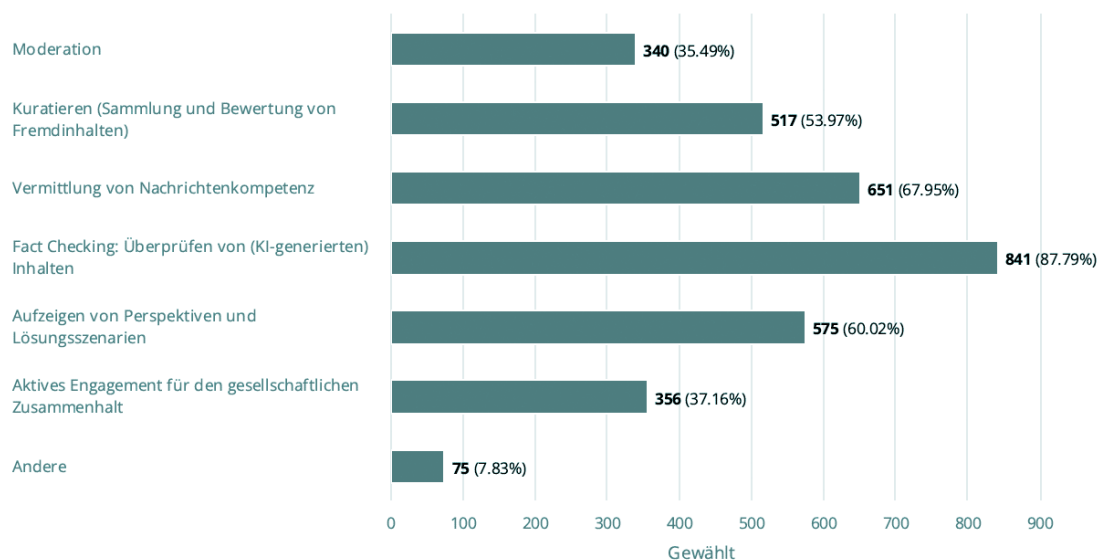


Abbildung 11: Mehrfachnennungen zu künftig stärker geforderten Rollen im Journalismus (in Prozent) (eigene Darstellung)

Diese Verschiebung vollzieht sich jedoch unter Bedingungen, die im Redaktionsalltag häufig als spannungsvoll erlebt werden, weil sich im Kolleg:innenkreis sehr unterschiedliche Geschwindigkeiten und Haltungen im Umgang mit digitaler Transformation und KI zeigen. Das ist an den Einschätzungen zur Veränderungsbereitschaft im Kolleg:innenkreis abzuleiten, bei der KI nicht explizit abgefragt wurde, aber wie alle Aspekte digitaler Transformation eine Rolle spielen dürfte. 63,0 Prozent der Journalist:innen unter 30 Jahren erkennen eine „eher hohe“ oder „sehr hohe“ generelle Veränderungsbereitschaft. In der Altersgruppe von 30 bis 50 Jahren liegt dieser Anteil bei 38,6 Prozent, bei den über 50-Jährigen sind es 23,1 Prozent. Umgekehrt bewerten 48 Prozent der Befragten die Veränderungsbereitschaft älterer Kolleg:innen als „eher gering“ oder „deutlich zu gering“ (n=965). Anpassung an KI-gestützte Arbeitsweisen wird damit als ungleiche Weiterentwicklung erlebt, die systemische Koordination erfordert und sich nicht allein über individuelle Lernbereitschaft der Mitarbeitenden auflösen lässt.

Die allgemeine Stimmung im Arbeitsumfeld verstärkt diesen Befund – auch wenn hier KI nicht explizit als Faktor abgefragt wurde. Über die Hälfte der Befragten (51,4 Prozent) bewerten die Stimmung in ihrem Arbeitsumfeld als „eher schlecht“, 20,1 Prozent als „miserabel“. Demgegenüber stehen 25,9 Prozent mit einer „eher guten“ und nur 2,6 Prozent mit

einer „ausgezeichneten“ Einschätzung (n=956) (Abb. 12). In den offenen Begründungen für eine schlechte oder miserable Stimmung wird KI punktuell genannt: 41 von 679 offenen Antworten (6,0 Prozent) beziehen sich explizit auf KI. Ähnlich fällt das Bild bei der Frage nach dem dringendsten Veränderungsbedarf im Arbeitsumfeld aus, bei der nur 14 von 934 Befragten (1,5 Prozent) KI ausdrücklich erwähnen. Diese geringen Anteile deuten darauf hin, dass *KI fast nie als isoliertes Problem* benannt wird. Sie drückt sich, so unsere Deutung, vielmehr in konkreten Belastungssituationen wie Zeitdruck, Ressourcenknappheit und fehlender struktureller Unterstützung aus.

Auf einer Skala von 1-4 (ausgezeichnet, eher gut, eher schlecht, miserable): Wie beurteilen Sie die allgemeine Stimmung in Ihrem Arbeitsumfeld?

Anzahl Antworten: 956

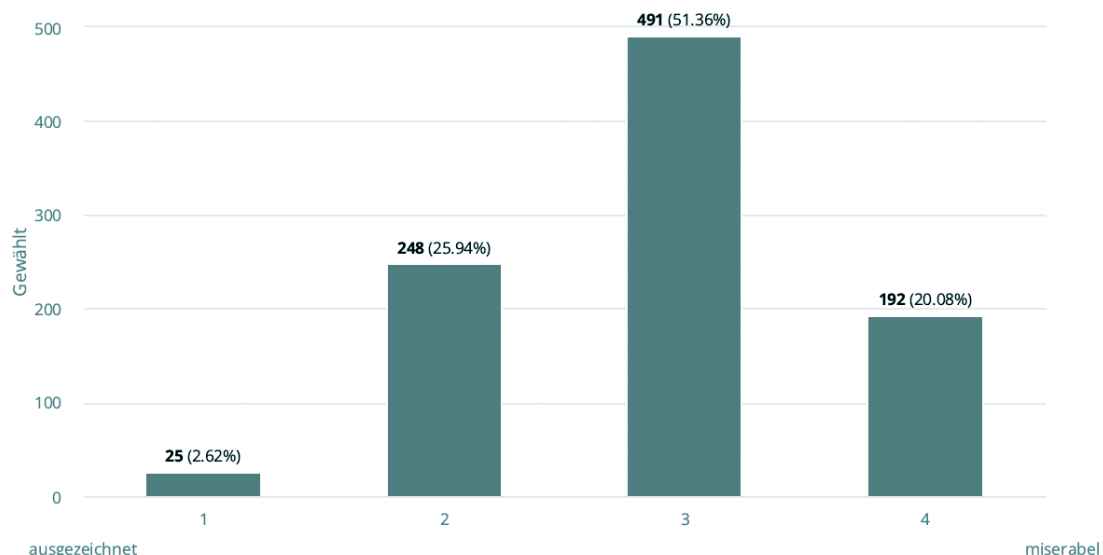


Abbildung 12: Einschätzung der allgemeinen Stimmung im eigenen Arbeitsumfeld durch befragte Journalist:innen (in Prozent) (eigene Darstellung)

Einen wichtigen Gegenpol zu diesen Belastungswahrnehmungen bildet die offene Frage nach dem Alleinstellungsmerkmal des Journalismus im Zeitalter von KI und Social Media (n=931). Die Antworten verdichten sich auf wenige wiederkehrende Motive: Einordnung, Kontextualisierung und Analyse; gründliche, häufig investigative Recherche; Verlässlichkeit und Glaubwürdigkeit; sowie die Fähigkeit zur Überprüfung und Verifikation von Informationen. Diese Selbstbeschreibungen markieren eine professionsbezogene Abgrenzung gegenüber automatisierten Informationssystemen wie KI-Chatbots, algorithmischen Newsfeeds oder automatisierter Contentproduktion. Journalismus wird dabei vor allem über nachvollziehbare Entscheidungsfindung definiert.

In der Zusammenschau zeichnen die Befragungsergebnisse ein konsistentes Bild, wie KI im journalistischen Alltag überwiegend als Faktor von der Branche wahrgenommen wird: Sie verschärft bestehende Unsicherheiten und macht neue Anforderungen sichtbar. Die hohe Sensibilität ihrer Risiken trifft zudem auf Arbeitsbedingungen, die eine organisationale und individuelle Resilienz erschweren. Leitlinien und normative Erwartungen sind zwar präsent, entfalten im Alltag jedoch nur begrenzte Entlastung: Der Umgang mit KI bleibt vielfach eine Frage persönlicher Bewältigung und situativer Anpassung.

Für die Frage der Resilienz ist dabei zentral, dass die geäußerten Sorgen auf ein Metier treffen, das bereits unter hohem Arbeitsdruck steht. In zahlreichen offenen Antworten werden Zeitmangel, Arbeitsverdichtung und fehlende Ressourcen als strukturelle Probleme benannt, die sich durch den Einsatz neuer Technologien offenbar weiter verschärfen. Die Befragung zeigt, dass der Journalismus in Deutschland gegenüber KI hoch sensibilisiert ist, Risiken klar benennt und zugleich an professionellen Maßstäben festhält. Gerade die Spannung zwischen ausgeprägter professioneller Normorientierung und begrenzten Handlungsspielräumen im redaktionellen Alltag bildet den empirischen Kern der Befunde.

Methodische Einordnung und analytische Reichweite der Befragung

Die vorliegenden Ergebnisse basieren auf einer quantitativen Befragung von Journalist:innen, die Mitglieder in Landesverbänden des DJV sind und sich damit hauptberuflich im Journalismus verorten. Die Befragung wurde als standardisierte Online-Erhebung unter mehr als 1.100 Mitgliedern des DJV durchgeführt. Sie fand im Zeitraum Ende Januar bis Anfang März 2025 statt und umfasste Journalist:innen unterschiedlicher Altersgruppen, Arbeitsfelder und Beschäftigungsformen. Ziel war nicht eine statistische Repräsentativität, sondern die Erfassung professioneller Einschätzungen zur Arbeitsrealität und zu den Auswirkungen von Digitalisierung und KI im Journalismus.

Gerade weil die Befragten aus dem journalistischen Feld selbst stammen, lassen sich die hier präsentierten Ergebnisse als Einblick in die beruflichen Wahrnehmungen und Erfahrungen von Journalist:innen im Umgang mit KI lesen. Sie geben Aufschluss darüber, wie KI im redaktionellen Alltag wahrgenommen, eingeordnet und bearbeitet wird – unabhängig davon, ob formale Leitlinien existieren oder regulatorische Vorgaben greifen. Die hohe Konsistenz der Antworten bei zentralen Risikoeinschätzungen spricht dafür, dass es sich um ähnliche Erfahrungsmuster handelt. Für den Argumentetransfer in dem vorliegenden Whitepaper sind die Befunde daher besonders relevant. Sie zeigen, dass normative Regelwerke und politische Regulierung auf eine Branche treffen, die bereits hoch sensibilisiert ist, zugleich aber unter strukturellen Bedingungen arbeitet, die Resilienz erschweren.

4.5 Leitlinien jenseits des Journalismus: Übertragbare Ansätze und blinde Flecken

Ein Blick über den Journalismus hinaus lässt erkennen, dass der Umgang mit KI in anderen gesellschaftlichen Bereichen und Berufsbranchen von ähnlichen Grundfragen geprägt ist, diese jedoch zum Teil völlig anders angegangen und beantwortet werden: Auch in Wissenschaft, Medizin, Verwaltung oder Bildung wird KI als Beschleuniger von Entscheidungen, Quelle neuer Abhängigkeiten und Herausforderung für Verantwortlichkeiten interpretiert. Der Unterschied liegt weniger in der Gegenwartsdiagnose als in der Reaktion: Während journalistische Leitlinien vor allem auf normative Selbstvergewisserung setzen, greifen andere Bereiche mitunter zu prozeduralen und organisatorischen Instrumenten, um technologische Unwägbarkeiten greifbar zu machen.

- In der **Wissenschaft** etwa steht nicht etwa Effizienz, sondern – ähnlich wie im Journalismus – Integrität im Mittelpunkt. KI wird dort als potenzieller Störfaktor epistemischer – also erkenntnisbezogener – Ordnung verstanden. Fragen der Urheberschaft, der Nachvollziehbarkeit und der Verantwortlichkeit werden deshalb klar adressiert, gerade weil wissenschaftliche Erkenntnisse auf Überprüfbarkeit und Reproduzierbarkeit angewiesen sind. KI darf deshalb unterstützen, aber nicht entscheiden; sie darf analysieren, aber nicht zuschreiben.⁹¹ Diese Grenzziehung soll die Bedingungen kollektiver Wissensproduktion schützen. Für den Journalismus ist diese Logik anschlussfähig, weil auch hier Glaubwürdigkeit und Vertrauen an Nachvollziehbarkeit (des Publikums) gebunden sind. Allerdings: Wissenschaft kann sich Zeit nehmen, Journalismus in der Regel kaum.
- Im **medizinischen Bereich** ist der Einsatz von KI von Beginn an als sicherheitsrelevante Intervention konzipiert. Leitlinien und Stellungnahmen – etwa zu klinischen Entscheidungsunterstützungssystemen – behandeln KI als Bestandteil klar definierter Behandlungsprozesse. Entsprechend detailliert sind die Regelungen zu Zuständigkeiten und Eingriffsmöglichkeiten: Es wird festgelegt, an welcher Stelle KI-Systeme Empfehlungen ausliefern dürfen, wer diese überprüft und unter welchen Bedingungen sie verwendet oder verworfen werden müssen. Fehler werden dabei nicht als Ausnahme verstanden, sondern als systemisch erwartbare Möglichkeit, auf die dann organisatorisch reagiert werden muss. Verantwortung wird deshalb nicht abstrakt zugeschrieben, sondern in Form von Verantwortungsketten organisiert, die von der Systementwicklung über den Einsatz bis zur letztverantwortlichen ärztlichen Entscheidung reichen.⁹² Die menschliche Entscheidungshoheit ist damit durch klar geregelte Abläufe praktisch abgesichert. Für den Journalismus ist vom medizinischen Bereich zu lernen, dass KI dort seit Jahren als Teil eines sicherheitskritischen Entscheidungszusammenhangs

verstanden wurde – also durchaus analog zu publizistischen Entscheidungen, die unter KI-Beteiligung öffentliche Orientierung, Vertrauen und demokratische Resilienz beeinflussen.

- Auch in der **öffentlichen Verwaltung** wird KI über verbindliche Verfahren eingebunden: Entscheidungen, die mithilfe von KI vorbereitet werden, müssen dokumentiert sein; Risiken werden vorab klassifiziert; und es wird definiert, welche Qualifikationen Mitarbeitende benötigen, um mit solchen Systemen arbeiten zu dürfen. KI gilt dabei nicht als neutrales Hilfsmittel, sondern als Instrument, das Verwaltungsentscheidungen beeinflussen und damit konkrete Auswirkungen auf Rechte, Ansprüche oder Lebenslagen von Bürger:innen haben kann. Verantwortung soll institutionell abgesichert werden – durch klare Zuständigkeiten und exakt überprüfbare Abläufe.⁹³ Für den Journalismus ist dieser Umgang allerdings schwer operationalisierbar, denn schließlich ist journalistisches Entscheiden sehr viel stärker situationsabhängig, zeitkritisch und kontextgebunden.
- Im **Bildungsbereich** wiederum steht weniger die Regulierung einzelner Anwendungen im Vordergrund als die Frage, welche Kompetenzen Lehrende und Lernende benötigen, um KI verantwortungsvoll einzusetzen. Leitlinien für Schule und Lehre verstehen KI daher weniger als Instrument denn als Anlass zur Reflexion professioneller Praxis: Wie funktionieren KI-gestützte Werkzeuge? Wo liegen ihre Grenzen? Und welche Folgen hat ihr Einsatz für pädagogische Beziehungen, schulische Bewertung und Lernprozesse per se? Transparenz gegenüber Lernenden, die Bewahrung pädagogischer Verantwortung und die eigene Ermächtigung gelten dabei als wichtige Voraussetzungen.⁹⁴ Diese Perspektive macht eine weitere Leerstelle im journalistischen Diskurs sichtbar: Dort wird in der Regel vorausgesetzt, dass Redaktionen über ausreichende Kompetenz im Umgang mit KI verfügen oder sich diese zumindest selbstständig aneignen können und lediglich deren Anwendung reguliert werden muss. Eine hieraus resultierende Leerstelle ist, dass, anders als im Bildungsbetrieb, der aktive und kontinuierliche Kompetenzerwerb der Medienschaffenden selbst im Bereich KI nur selten priorisiert wird.

Gemeinsam ist diesen nicht-journalistischen Leitlinien eine allgemeine *Skepsis gegenüber der vollständigen Kontrolle von KI*. Statt Kontrolle zu versprechen, organisieren sie Verfahren, mit denen ihre Wirkungen begrenzt, überprüft und im Bedarfsfall korrigiert werden können. Resilienz entsteht hier nicht aus einem Grundvertrauen in die neue Technologie, sondern aus *bewusster Organisation von Unsicherheit*. Für den Journalismus ist diese Haltung besonders relevant, weil er unter öffentlichem Druck (re)agiert und seine publizistische Verantwortung, kritische Öffentlichkeit herzustellen, nicht ohne Weiteres delegieren kann, also auch sein alltägliches ‚Chaos‘ zwangsläufig organisieren muss.

Der Vergleich macht deutlich: Journalismus kann von anderen Bereichen Wichtiges lernen, teilweise adaptieren. Leitlinien jenseits des Journalismus zeigen außerdem, dass Resilienz dort entsteht, wo Werte und Ansprüche in möglichst konkrete Entscheidungs- und Organisationsstrukturen übersetzt werden. Genau an dieser Transferleistung hapert es im Journalismus bislang.

4.6 Zwischenfazit: Von normativer Selbstbindung zu organisationaler Handlungsfähigkeit

Die überblicksartige Zusammenstellung von Leitlinien, Regelwerken und Positionspapieren zeigt, dass der Journalismus auf den Einsatz von KI keineswegs unvorbereitet oder willkürlich reagieren muss. Im Gegenteil: Die Vielzahl der ethischen Wegmarken im Jahr 2026 ist vielmehr das Signal einer bemerkenswerten normativen Basis. Gleichzeitig wird in der Zusammenschau deutlich, dass diese Selbstvergewisserung an unterschiedliche Grenzen stößt. Leitlinien entfalten ihre Wirkung primär nach innen: Sie strukturieren Erwartungen, stabilisieren professionelle Identität und etablieren Grenzen. Was sie jedoch nicht leisten, ist die Bearbeitung jener Abhängigkeiten, die außerhalb redaktioneller Entscheidungsmacht liegen: Datenhoheit und -sicherheit, Trainingsgrundlagen von KI-Modellen, Plattformlogiken, Markt- und Machtkonzentration, ökologische Nachhaltigkeitsfragen im Zusammenhang mit KI-Serverfarmen sowie die rechtliche Durchsetzung entziehen sich weitgehend journalistischer Selbstregulierung.

Gerade hier zeigt sich jedoch eine paradoxe Dynamik: Je klarer und konziser journalistische Leitlinien formuliert sind, desto erkennbarer werden auch die Bedingungen, denen sie nicht genügen (können). Denn selbst dort, wo Redaktionen journalistische Verantwortung übernehmen und Transparenz versprechen, bleiben sie eingebettet in technologische und ökonomische Zwänge, auf die sie selbst keinen oder nur bedingt Einfluss nehmen können. Resilienz lässt sich unter diesen Bedingungen deshalb nicht als rein organisationale Leistung denken. Sie verweist auf eine Ebene, bei der die journalistische Selbstverpflichtung auf eine politisch-gesellschaftliche Rahmensetzung trifft.

Damit verschiebt sich der Fokus von dem, was Journalismus intern regeln kann, zu dem, welche externen Bedingungen erforderlich sind, um diese Selbstbindung wirksam werden zu lassen. Resilienz im KI-Zeitalter erweist sich damit als klar geteilte Verantwortung. Sie verlangt nicht nur professionelle Haltung und ethische Selbstverortung, sondern auch rechtliche, infrastrukturelle und medienpolitische Voraussetzungen, die journalistische Autonomie sichern helfen.

5. Medienpolitik und Regulierung: Zwischen Förderung und Kontrollillusion

Worauf dieses Kapitel Antworten gibt:

1. Warum Regulierung im KI-Kontext dort ansetzen muss, wo freiwillige Selbstbindung an ihre Grenzen stößt.
2. Welche Logiken, Ziele und Risikokategorien der europäische AI Act verfolgt.
3. Wie journalistische Praxis im AI Act adressiert wird – und wo Unschärfen und Leerstellen bleiben.
4. Warum Public-Interest-KI als gemeinwohlorientierte Alternative diskutiert werden muss – und wo sie praktisch scheitert.
5. Welche Positionen Verbände und Akteur:innen im Journalismus zur Regulierung einnehmen.
6. Warum Regulierung Digitale Resilienz nicht ersetzt, sondern neue Anforderungen an redaktionelle Organisationen stellt.

5.1 Von redaktioneller Selbstbindung zu öffentlicher Rahmensetzung

Als im Oktober 2025 die Europäische Rundfunkunion (EBU) die Ergebnisse ihrer groß angelegten Studie „News Integrity in AI Assistants“⁹⁵ präsentierte, an der auch ARD und ZDF beteiligt waren, offenbarte sich ein alarmierendes Bild: 45 Prozent aller KI-Antworten zu nachrichtlichen Inhalten wiesen mindestens einen erheblichen Fehler auf – unabhängig von Sprache, Region oder Plattform. Besonders deutlich wurde das Problem an einem einfachen Beispiel: Mehrere KI-Assistenten bezeichneten weiterhin Olaf Scholz als deutschen Bundeskanzler, obwohl Friedrich Merz das Amt bereits einen Monat zuvor übernommen hatte. Der Fall machte sichtbar, dass KI-Systeme aktuelle politische Informationen häufig unzuverlässig wiedergeben, zugleich aber mit dem Anspruch autoritativer Auskunft auftreten. Genau diese Kombination aus Reichweite, Fehleranfälligkeit und fehlender Verantwortungsstruktur löste erneute Appelle an die Tech-Konzerne aus, sich auf verbindliche Austauschformate einzulassen, mehr Transparenz über Datenquellen herzustellen und journalistische Qualitätsangebote stärker zu berücksichtigen.

Diese und viele andere dokumentierten Fehler deuten nicht auf ein technisches Versagen einzelner Modelle hin, sondern auf die Risiken, die mit der Art der Nutzung dieser Technologien einhergehen: Wenn – wie die EBU-Studie ebenfalls ermittelte – sieben Prozent der Online-Nachrichtenkonsument:innen KI-Assistenten nutzen, um Nachrichten abzurufen (bei den unter 25-Jährigen bereits 15 Prozent), wird die systematische Fehlerhaftigkeit generativer KI bei unzureichender Nutzungskompetenz zur demokratischen Gretchenfrage.

Das vorangegangene Kapitel hat gezeigt, dass der Journalismus auf die Herausforderungen von KI mit normativer Selbstbindung reagiert. Dies ist der Versuch, technologische Dynamiken an journalistische Grundwerte zurückzubinden, um sie zu schützen. Selbst dort, wo Redaktionen diese Verantwortung übernehmen und Transparenz herstellen wollen, stoßen sie schnell an Grenzen. Die Frage nach einer **KI-Resilienz** verlagert sich deshalb zwangsläufig von der Ebene interner Selbstregulierung auf die Ebene öffentlicher Rahmensetzung, die in diesem Abschnitt diskutiert werden.

Vor diesem Hintergrund tritt eine Re-Orientierung der journalistischen Leistung von einer Aufmerksamkeitsökonomie (Plattformlogik, Reichweite) hin zu einer ‚Vertrauensökonomie‘ (Beziehungsarbeit, Glaubwürdigkeit) (Abb. 13) in den Vordergrund. Dieser Wandel wird von Regulierungsmaßnahmen wie dem europäischen AI Act gefördert: Journalismus kann sich dadurch als Infrastruktur jenseits reiner datenökonomisch getriebener Logiken neu profilieren und möglicherweise stabilisieren.



Abbildung 13: Übergang von der Aufmerksamkeitsökonomie zur Vertrauensökonomie – Journalismus als vermittelnde Infrastruktur im KI-Zeitalter (eigene Darstellung)

Medienpolitik und Regulierung treten vor diesem Hintergrund weniger als nachgelagerte Kontroll- denn als *konstitutive Ermöglichungsinstanzen journalistischer Handlungsfähigkeit* auf. Regulierung wird hier deshalb nicht als Eingriff in redaktionelle Autonomie verstanden, sondern als Bedingung dafür, dass diese unter veränderten technologischen Bedingungen überhaupt erhalten bleibt: Wo Märkte hochkonzentriert sind, technische Systeme intransparent agieren und Verantwortlichkeiten diffuser werden, kann journalistische Selbstbindung ihre stabilisierende Wirkung nur entfalten, wenn sie dauerhaft durch geeignete rechtliche und politische Rahmenbedingungen gestützt wird.

Resilienz im Journalismus erscheint damit als geteilte Aufgabe. Sie verlangt eine reflektierte Einstellung gegenüber Einsatzmöglichkeiten neuer Technologien in der individuellen und organisationalen Praxis. Und zugleich erfordert sie eine sinnvolle Regulierung, die digitale Machtasymmetrien adressiert, Transparenz ermöglicht und Schutzräume für unabhängige Berichterstattung sichert (vgl. Kap. 3). Diese Einstellung macht sich also seitens der Redaktionen durch eine kritische Distanz bei gleichzeitiger Innovationsoffenheit bemerkbar. Diese Perspektive bildet auch den gedanklichen Ausgangspunkt für die folgende Betrachtung der regulatorischen Ebenen, Zuständigkeiten und Instrumente, die den Einsatz von KI im Journalismus durchgehend prägen (sollten).

5.2 Der EU AI Act und der Journalismus: Rahmenbedingungen statt Medienregulierung

Mit dem EU AI Act hat die Europäische Union erstmals einen umfassenden Rechtsrahmen für künstliche Intelligenz geschaffen.⁹⁶ Für den Journalismus ist diese Regulierung weniger deshalb relevant, weil sie redaktionelle Arbeit direkt betrifft, sondern weil sie die technischen Systeme reguliert, auf denen Journalismus zunehmend aufbaut.

Der AI Act verfolgt einen risikobasierten Ansatz: KI-Systeme werden je nach möglichem Schaden für Grundrechte, Sicherheit oder gesellschaftliche Prozesse unterschiedlich streng reguliert. Journalistische Anwendungen werden dabei nicht als eigener Bereich behandelt. Das ist bewusst so angelegt, um Presse- und Meinungsfreiheit nicht durch staatliche Aufsicht über redaktionelle Prozesse zu gefährden.

Dennoch hat der AI Act mehrere indirekte Auswirkungen auf journalistische Arbeit: Erstens schafft er **Transparenzpflichten für generative KI-Systeme**. Anbieter großer KI-Modelle müssen künftig offenlegen, wie ihre Systeme funktionieren, welche Sicherheitsmaßnahmen sie nutzen und welche urheberrechtlich geschützten Inhalte möglicherweise in Trainingsdaten enthalten sind. Für Medienhäuser ist das relevant, weil ihre Inhalte zu den wichtigsten Trainingsquellen vieler KI-Systeme gehören.

Zweitens stärkt der AI Act **Kennzeichnungspflichten für KI-generierte Inhalte**. Wenn Inhalte vollständig durch KI erzeugt wurden oder synthetische Medien wie Deepfakes entstehen, müssen diese gekennzeichnet werden. Diese Vorgaben unterstützen journalistische Bemühungen, zwischen verifizierten Informationen und automatisierten Inhalten zu unterscheiden.

Drittens setzt der AI Act auf die **Regulierung einflussreicher Basismodelle (General Purpose AI)**. Anbieter solcher Modelle müssen Risiken bewerten, Missbrauch verhindern und technische Dokumentationen bereitstellen. Da viele redaktionelle KI-Anwendungen auf solchen Modellen aufbauen, betrifft diese Regulierung indirekt auch journalistische Arbeitsprozesse.

Gleichzeitig lässt der AI Act zentrale Fragen offen, die für den Journalismus zunehmend relevant werden: Unklar bleibt etwa, wie mit der **Nutzung journalistischer Inhalte als Trainingsdaten** umzugehen ist. Zwar verlangt der AI Act mehr Transparenz über Datenquellen, er schafft jedoch keinen eigenständigen Schutzmechanismus für journalistische Inhalte. Medienhäuser müssen ihre Ansprüche weiterhin vor allem über Urheberrecht oder Lizenzmodelle durchsetzen. Auch **hybride Formen der Zusammenarbeit zwischen Mensch und Maschine** werden regulatorisch kaum adressiert. Wenn KI-Systeme Recherche unterstützen, Texte strukturieren oder Zusammenfassungen erstellen, bleibt die Verantwortung rechtlich bei den Menschen, während die technische Vorstrukturierung journalistischer Entscheidungen weiter zunimmt.

Der AI Act reguliert damit vor allem die **Infrastruktur der KI-Systeme**, nicht die Praxis des Journalismus selbst. Für Redaktionen bedeutet das zweierlei: Einerseits schützt die Regulierung vor intransparenten und potenziell manipulativen KI-Systemen. Andererseits bleibt es Aufgabe der Medienorganisationen, diese allgemeinen Vorgaben in konkrete redaktionelle Regeln und Arbeitsabläufe zu übersetzen.

5.3 Gegenentwurf zu Big Tech: Public-Interest-KI als gemeinwohlorientierte Ermöglichungsstrategie

Der AI Act adressiert Risiken, schafft aber durch grundrechtliche Zurückhaltung eine Resilienz-Lücke. Regulierung definiert, was journalistische KI nicht sein darf – manipulativ, intransparent, diskriminierend. Was sie jedoch sein könnte und wer sie unter welchen Bedingungen bereitstellt, bleibt ungeklärt.

Redaktionen befinden sich somit in einer grundlegenden Abhängigkeit von proprietären Infrastrukturen weniger Technologiekonzerne: Erstens haben Plattformen ihre Modelle

auf journalistischen Inhalten trainiert, ohne Lizenzen zu erwerben oder Transparenz herzustellen.⁹⁷ Verlage und Sender wiederum investieren seit jeher in Recherche, Verifikation und redaktionelle Qualität – die Tech-Konzerne verwerten diese Arbeit, ohne dass deren Urheber:innen Kontrolle oder Einblick hätten. Zweitens fließen schätzungsweise 83 Prozent der gesamten jährlichen europäischen Cloud- und Software-Ausgaben an US-amerikanische Anbieter, wie eine Studie des Unternehmensverbands Cigref belegt.⁹⁸ Das betrifft auch Medienorganisationen. Drittens binden proprietäre Schnittstellen, intransparente Algorithmen und einseitige Preisgestaltung die Redaktionen an einzelne Anbieter – Verhandlungs- oder Gestaltungsmacht existiert so gut wie nicht.

Parallel zur Regulierungsdebatte hat sich deshalb eine konzeptionelle, diskursive und entwicklungspolitische Bewegung formiert, die einen Gegenentwurf verfolgt: Public-Interest-KI als *infrastrukturelle Resilienz-Strategie*. Zwar ist dieser Diskurs begrifflich noch recht diffus, normativ überfrachtet und empirisch fragmentiert. Dennoch ist das zugrundeliegende Vorhaben hilfreich, um zu verstehen, ob und wie gemeinwohlorientierte KI-Infrastrukturen tatsächlich zur journalistischen Widerstandsfähigkeit beitragen können.

Begriffliche Unschärfen: Was meint „im öffentlichen Interesse“?

In der aktuellen Debatte über die gesellschaftliche Gestaltung von KI-Technologien hat sich der Begriff „Public Interest AI“ als analytischer Bezugspunkt etabliert.⁹⁹ Der Ansatz richtet den Blick darauf, inwiefern Entwicklung, Einsatz und Regulierung von KI zum langfristigen Wohlergehen der Gesellschaft beitragen und zentrale öffentliche Güter wie demokratische Entscheidungsprozesse, informationelle Selbstbestimmung oder gesellschaftliche Teilhabe stärken. KI erscheint in dieser Perspektive als Teil gesellschaftlicher Infrastruktur, deren Gestaltung am Gemeinwohl orientiert sein soll.¹⁰⁰ Zugleich gilt der Begriff des öffentlichen Interesses als normativ offen und kontextabhängig; denn was jeweils als gemeinwohlorientiert gilt, entsteht in politischen und gesellschaftlichen Aushandlungsprozessen und verändert sich mit den jeweiligen gesellschaftlichen Rahmenbedingungen.

Für KI-Infrastrukturen im Journalismus ist diese begriffliche Unschärfe nicht nur theoretisch relevant – sie hat praktische Konsequenzen: Wenn unklar bleibt, wer *öffentliches Interesse* wie definiert, bleibt auch fraglich, wer Public-Interest-KI kontrolliert, finanziert und steuert. Sollen es Regierungen sein? Dann droht staatliche Einflussnahme auf journalistische Werkzeuge – ein Widerspruch zur Pressefreiheit. Sollen es Tech-Unternehmen sein, die sich als *gemeinwohlorientiert* bezeichnen? Dann bleiben Profit- und Skalierungslogiken bestehen. Sollen es zivilgesellschaftliche Multi-Stakeholder-Gremien sein? Dann stellt sich die Frage, wie Entscheidungsfähigkeit bei divergierenden Interessen

gesichert wird. Die begriffliche Offenheit wird für journalistische Resilienz zum strukturellen Problem: Redaktionen benötigen verlässliche, unabhängige Infrastrukturen, nicht zwingend weitere Aushandlungsarenen.

Anders akzentuiert ist *Public AI*, wie das Fraunhofer-Institut für System- und Innovationsforschung den Begriff verwendet.¹⁰¹ Hier rücken KI-Infrastrukturen in den Vordergrund, die unter öffentlicher Beteiligung oder Finanzierung entstehen. Unterschieden werden hier drei Dimensionen: *Vertrauenswürdigkeit* (Datenschutz, Fairness, Transparenz), *soziale Innovation* (Fokus auf gesellschaftliche Herausforderungen statt Profitmaximierung) und *KI als Gemeingut* (offener Zugang, partizipative Entwicklung).

Für journalistische Resilienz ist entscheidend, dass diese Dimensionen nicht automatisch erfüllt sind, nur weil etwas öffentlich finanziert wird. Auch staatlich getragene KI-Systeme können intransparent operieren, diskriminieren oder politischen Übergriffen ausgesetzt sein. So wurde das niederländische Sozialbetrugserkennungssystem SyRI als öffentlicher Dienst verkauft, diskriminierte aber systematisch und verweigerte Transparenz über seine Funktionsweise – bis Gerichte es stoppten.¹⁰² Übertragen auf den Journalismus bedeutet das: Selbst wenn öffentlich-rechtliche Sender oder Medienanstalten KI-Tools entwickelten oder finanzierten, garantiere das im Zweifel weder redaktionelle Unabhängigkeit noch Schutz vor technokratischer Steuerung. Die Frage ist nicht, ob Infrastrukturen öffentlich finanziert werden, sondern wie ihre Steuerung gestaltet ist – und ob diese journalistische Autonomie tatsächlich strukturell absichern kann.

Eine dritte begriffliche Variante kreist um *Open Source AI*. Initiativen wie LAION argumentieren, dass frei verfügbarer Code Transparenz ermöglicht, Wettbewerb fördert und digitale Souveränität sichert.¹⁰³ Doch der Begriff ist mittlerweile umkämpft. Meta bewirbt sein Modell Llama 2 als „Open Source“, obwohl die Lizenz restriktiv und kommerzielle Nutzung beschränkt ist.¹⁰⁴ Selbst die Open Source Initiative hat 2024 eine Definition vorgelegt, die keine Offenlegung der Trainingsdaten verlangt. Für Redaktionen heißt das: Ohne Einblick in die Datengrundlagen können sie nicht prüfen, welche Inhalte (etwa ihre eigenen) zum Training verwendet wurden. „Open Source“ wird auf diese Weise zur Legitimationsfassade, hinter der sich dieselbe Intransparenz verbirgt wie bei proprietären Systemen. Diese begriffliche Unschärfe ist kein akademisches Glasperlenspiel. Sie verdeckt, dass drei analytisch trennbare Dimensionen vermischt werden:

- **Governance:** Wer kontrolliert die KI?
- **Infrastruktur:** Wer stellt sie bereit?
- **Transparenz:** Wer kann sie prüfen?

Nur wenn alle drei Dimensionen übereinandergelegt werden, könnte Public-Interest-KI tatsächlich zur journalistischen Resilienz beitragen. Governance-Maßnahmen allein –

etwa durch Ethik-Beiräte bei Tech-Konzernen – reichen nicht, solange die Infrastruktur proprietär bleibt. Infrastruktur allein – etwa staatliche Cloud-Dienste – genügt nicht, solange die Governance intransparent ist. Und Transparenz allein – etwa durch öffentlich einsehbaren Code – hilft nicht, wenn die Nutzung der Trainingsdaten im Dunkeln bleibt.

Trainingsdaten: Die ungelöste Kernfrage

Der AI Act verpflichtet Anbieter generativer KI zu mehr Transparenz über ihre Trainingsdaten.¹⁰⁵ Zugleich sieht die Verordnung jedoch vor, dass Geschäftsgeheimnisse geschützt bleiben – eine Klausel, die die Transparenzpflicht faktisch relativiert. Für den Journalismus entsteht ein doppeltes Resilienz-Problem: Einerseits wurde journalistische Arbeit bereits über viele Jahre beim Training von großen Sprachmodellen ausgebeutet, andererseits reproduzieren KI-Assistenten diese Inhalte fehlerhaft und dekontextualisiert.

Parallel entstehen erste Initiativen, die dieses Ungleichgewicht adressieren sollen. Programme wie „Current AI“, eine 2025 in Paris gestartete Public-Interest-AI-Initiative mit einem Startvolumen von rund 400 Millionen US-Dollar, arbeiten an Modellen für kollektive Lizenzierung und Vergütung journalistischer Inhalte.¹⁰⁶ Im Programmbereich „Public Interest Media“ soll eine Infrastruktur entstehen, über die Medienorganisationen ihre Inhalte gebündelt lizenzieren und gegenüber KI-Anbietern gemeinsam verhandeln können.

Parallel dazu wird über europäische „Data Commons“¹⁰⁷ nachgedacht: kollektive Datenverwaltung unter öffentlicher Kontrolle, die alternative Trainingsmodelle ermöglichen soll. Solche Strukturen könnten Trainingsdaten für KI bereitstellen, die Urheberrechte respektieren und gleichzeitig einen offenen Zugang für Forschung und gemeinwohlorientierte Anwendungen ebnen.

Diese Ansätze reagieren auf eine Entwicklung, die in der KI-Governance auch als „Data Winter“¹⁰⁸ beschrieben wird: Der Zugang zu hochwertigen Trainingsdaten konzentriert sich bei wenigen Technologieunternehmen, während Informationen aus offenen Quellen systematisch abgeschöpft werden. Medienorganisationen verlieren damit die Kontrolle über ihre Inhalte als Trainingsressource, erhalten jedoch kaum Einblick in die entstehenden Modelle.

Weder kollektive Lizenzmodelle noch Data-Commons-Konzepte haben bislang eine operative Infrastruktur hervorgebracht. Offen bleibt daher eine zentrale Governance-Frage: Wer kontrolliert künftig die Datenräume, auf denen KI-Systeme trainiert werden – staatliche Institutionen, Branchenallianzen oder zivilgesellschaftliche Treuhandmodelle? Für den Journalismus entscheidet sich an dieser Frage, ob redaktionelle Inhalte künftig als

frei verfügbare Rohstoffe digitaler Plattformen fungieren oder als Teil einer öffentlich verantworteten Informationsinfrastruktur behandelt werden.

Kritiker:innen weisen darüber hinaus auf fundamentale Transparenzdefizite hin: Es fehle an klarer Dokumentation der Datenherkunft, an Informationen über Erhebungsprozesse und an Nachweisen der Einwilligung abgebildeter Personen oder Urheber:innen.¹⁰⁹ Zugleich trainiert LAION auf Material aus Common Crawl, das systematisch urheberrechtlich geschützte Inhalte enthält. Das heißt: Verlage haben keine Kontrolle darüber, ob und wie ihre Inhalte verwendet werden – und geizen nicht mit Kritik.¹¹⁰ Der Code mag offen sein, die Daten bleiben intransparent. Die Mozilla Foundation und der Think Tank Open Future kritisieren zudem, dass der AI Act Open-Source-Modelle privilegiert und ihnen weniger strenge Transparenzpflichten auferlegt als proprietären Systemen.¹¹¹ Dieser „Frankenstein-Ansatz“, wie es bei Open Future heißt,¹¹² ermögliche es, dass vermeintlich offene Modelle faktisch weniger transparent und weniger gut dokumentiert sind als geschlossene. Für die Resilienz des Journalismus ist das kontraproduktiv: Ohne verpflichtende Trainingsdaten-Transparenz – einschließlich Herkunft, Lizenzierung und Vergütung – bleibt weder proprietäre noch Open-Source-KI verantwortbar.

Recheninfrastruktur: Souveränität als Illusion?

Selbst wenn Redaktionen vermeintlich offene Modelle nutzen, laufen diese auf geschlossenen Cloud-Infrastrukturen. Dies erzeugt dreifache Abhängigkeit: ökonomisch durch explodierende Kosten und einseitige Preisgestaltung, rechtlich durch Datenzugriffe nach US-amerikanischem Recht, technisch durch sogenannten ‚Vendor Lock-in‘. Gemeint ist damit eine Abhängigkeitsstruktur, in der Kund:innen durch technische und vertragliche Bindungen an einen Anbieter gekoppelt bleiben und ein Wechsel nur mit erheblichem Aufwand möglich ist. Migrationen zu anderen Systemen sind meist kostspielig, komplex und mit technischen Risiken verbunden.

GAIA-X sollte hier Abhilfe schaffen. Das europäische Cloud-Projekt versprach eine Infrastruktur mit hohen Datenschutzstandards als Alternative zu Angeboten großer Cloud-Anbieter wie Amazon Web Services oder Microsoft Azure. Die erhoffte Marktdurchdringung blieb jedoch aus. Eine offizielle Evaluationsstudie bescheinigt zwar Fortschritte bei der technischen Umsetzung, verweist aber zugleich auf ungelöste Finanzierungsfragen und eine geringe Nutzerorientierung.¹¹³ Kritiker:innen sprechen von einem „Paper Monster“¹¹⁴ ohne greifbare Ergebnisse und sehen die ursprünglichen Ziele durch die Einbindung großer Cloud-Konzerne unterlaufen. Hinzu kommt ein strukturelles Problem: GAIA-X stellt vor allem Rechenkapazität bereit, aber kaum anwendungsnahe Werkzeuge für den journalistischen Alltag. Redaktionen benötigen keine abstrakte Infrastruktur, sondern

einsatzfähige Anwendungen – etwa für Transkription, Übersetzung oder Verifikation. Auch die politische Debatte über europäische Supercomputer zielt bislang vor allem auf wissenschaftliche Forschung. Für Redaktionen sind jedoch andere Anforderungen entscheidend: stabile Systeme im Dauerbetrieb statt zeitlich begrenzter Rechenkontingente für Experimente. Rechenhoheit allein reicht deshalb nicht aus, um journalistische Resilienz zu stärken. Selbst mit europäischen Rechenzentren bleibt offen, wer die Modelle entwickelt, betreibt und weiterentwickelt – und nach welchen Kriterien dies geschieht.

KI-Governance: Die verdrängte Machtfrage

Das regulatorische Dilemma, das wir in Kapitel 5.2 beschrieben haben, wiederholt sich bei der Governance gemeinwohlorientierter KI: Strikte staatliche Kontrolle könnte die Pressefreiheit gefährden, während privatwirtschaftliche Strukturen – selbst wenn sie sich als gemeinwohlorientiert präsentieren – Exit-Druck und Skalierungszwängen unterliegen. Open-Source-Communities wiederum sind oft nicht inklusiv, sondern von Tech-Eliten dominiert, denen journalistische Expertise fehlt.

Das Forum on Information and Democracy hat ein freiwilliges Zertifizierungssystem nach dem Vorbild von Fairtrade vorgeschlagen:¹¹⁵ Multi-Stakeholder-Gremien mit mindestens 50 Prozent Beteiligung aus dem globalen Süden sollen Standards setzen und unabhängige Prüfstellen deren Einhaltung kontrollieren. Doch das System bleibt anfällig: Wer finanziert diese Prüfstellen? Wer setzt Standards durch, wenn Anbieter sich verweigern?

Der Vorschlag mag Orientierung bieten, strukturelle Machtverhältnisse kann er wahrscheinlich nicht verändern. Wikimedia und die Gewerkschaft Erziehung und Wissenschaft fordern wiederum eine unabhängige, staatlich finanzierte Prüfstelle für KI in Bildung und Medien.¹¹⁶ Doch wie lässt sich deren Unabhängigkeit sichern, wenn das Geld vom Staat kommt? Soll die Prüfstelle nach dem Presserat-Modell funktionieren, also als Selbstregulierung der Branche? Oder nach dem Datenschutzbehörden-Modell als staatliche Aufsicht? Für journalistische Resilienz wäre entscheidend, dass redaktionelle Autonomie auch bei gemeinwohlorientierter KI gewahrt bleibt.

Empirische Realität: Viel Diskurs, wenig Infrastruktur

Public-Interest-AI-Initiativen nutzen den KI-Boom durchaus strategisch, um Fördergelder einzuwerben – doch hier liegt eine strukturelle Gefahr: Gesellschaftliche Probleme werden dabei tendenziell technologisch umgedeutet. Die Vertrauenskrise im Journalismus erscheint dann nicht mehr als Frage nach tragfähigen Geschäftsmodellen oder

redaktioneller Qualität, sondern schlicht als ein ‚KI-Tool-Problem‘, das sich mit dem richtigen Werkzeug lösen lässt. Was dabei systematisch ausgeblendet wird, ist der Zeithorizont: Projektbasierte Finanzierung bemisst sich in Monaten, belastbare Institutionen aber wachsen über Jahre.¹¹⁷ Eine kartografische Auswertung von Public-Interest-AI-Projekten zeigt entsprechend, dass viele von ihnen bereits inaktiv sind – sie haben die Förderphase nicht überlebt.¹¹⁸ Für Redaktionen, die auf solche Werkzeuge setzen, entstehen daraus neue Abhängigkeiten: Verschwinden die Tools, müssen Workflows von Grund auf neu aufgebaut werden.

Ein aktuelles Beispiel illustriert das strukturelle Machtungleichgewicht: Die Europäische Union hat im Juli 2025 einen „Code of Practice“ für generative KI veröffentlicht,¹¹⁹ der den AI Act konkretisieren soll. Der Code ist rechtlich nicht bindend – große Tech-Konzerne wie Amazon, Microsoft, OpenAI und Anthropic haben unterzeichnet, andere wie Meta lehnen dessen Unterzeichnung ab und setzen auf eigene Compliance-Strategien. Der AI Act lässt diese Wahlfreiheit zwar explizit zu (Art. 56), dies schwächt aber die Durchsetzbarkeit: Anbieter können den Code umgehen, indem sie eigene Maßnahmen nachweisen – ohne sich einer externen Prüfung unterziehen zu müssen.

Hinzu kommt: Während systemische Risiken wie Cyberangriffe explizit geregelt werden, bleiben Maßnahmen gegen Desinformation – das eigentliche Kernproblem für den Journalismus – weitgehend freiwillig. Im Dezember 2025 legte die EU-Kommission einen Entwurf für Transparenzregeln vor. Das Ziel: KI-generierte Inhalte sollen für Nutzer:innen erkennbar sein – durch digitale Wasserzeichen, Metadaten oder Kennzeichnungen wie „erstellt mit KI“.¹²⁰ Doch es gibt einen Haken: Diese Regeln zielen nur auf die Herkunft eines Inhalts, nicht auf seinen Wahrheitsgehalt. Ein Interview, das vollständig erfunden ist – also Aussagen enthält, die eine Person nie gemacht hat –, bleibt Desinformation, auch wenn es korrekt als „KI-generiert“ gekennzeichnet ist. Das Label beschreibt die Produktionsweise, nicht die Faktentreue. Für den Journalismus löst das die Vertrauenskrise nicht.

Diese Ausgangslage macht es für gemeinwohlorientierte Alternativen doppelt schwer: Zum einen konkurrieren sie mit Anbietern, die über Größenvorteile verfügen – Rechenkapazität, Trainingsdaten und Entwicklungsressourcen. Zum anderen fehlt ihnen die institutionelle Verankerung: Während proprietäre Modelle in cloud-basierte Ökosysteme eingebettet sind (AWS, Azure, Google Cloud), müssen offene Modelle jedes Mal neu in bestehende Workflows integriert werden. Projektbasierte Förderung – sei es durch staatliche oder zivilgesellschaftliche Drittmittel – verschärft das Problem: Nach zwei oder drei Jahren endet die Finanzierung, doch die Wartung von Modellen, Datensätzen und APIs läuft weiter. Ohne langfristige institutionelle Trägerschaft entstehen keine belastbaren

Alternativen. Das Ergebnis: Redaktionen setzen weiterhin auf kommerzielle Anbieter, nicht aus Überzeugung, sondern aus Notwendigkeit. Die Abhängigkeiten verfestigen sich.

Resilienz-Potenzial und strukturelle Grenzen

Was bedeutet das für journalistische Resilienz?

Erstens ist **begriffliche Präzisierung notwendig**. Der dominante Diskurs vermischt Governance, Infrastruktur und Transparenz. Für Resilienz im Journalismus braucht es jedoch konzeptionelle Klarheit: Public-Interest-KI trägt nur dann zur Widerstandsfähigkeit bei, wenn alle drei Dimensionen erfüllt sind.

Zweitens: **Regulierung und Public-Interest-KI sollten sich ergänzen** – tun es aber nicht gleichgewichtig. Der AI Act setzt Grenzen für eine problematische KI-Nutzung, schafft jedoch keine Alternativen: Wenn nur proprietäre Tools verfügbar sind, können Redaktionen nicht auf offene Modelle ausweichen. Gemeinwohlorientierte KI-Infrastrukturen könnten solche Alternativen bieten. Doch Governance, Nachhaltigkeit und Skalierung sind ungeklärt. Diese Asymmetrie ist entscheidend: Regulierung ist verbindlich, und wer gegen sie verstößt, muss mit Sanktionen rechnen. Public-Interest-KI hingegen bleibt freiwillig und abhängig von Förderpolitik, Projektfinanzierung und gutem Willen.

Drittens: **Was gemeinwohlorientierte offene Modelle zur KI-Resilienz im Journalismus beitragen könnten, bleibt Theorie**. Würde Public-Interest-KI tatsächlich Autonomie über Trainingsdaten, Unabhängigkeit von proprietären Recheninfrastrukturen und partizipative Governance ermöglichen, wäre der Resilienz-Gewinn erheblich. Faktisch aber bleibt die Trainingsdaten-Governance ungeklärt, die Recheninfrastruktur fragmentiert und unterfinanziert, die Governance begrifflich diffus und institutionell ungelöst. Nachhaltigkeit scheitert an Projektförderung und Hype-Zyklen – dauerhafte Strukturen entstehen auf diese Weise nicht.

Public-Interest-KI sollte Redaktionen entlasten, tatsächlich erzeugt sie neue Komplexität. Statt nur den AI Act in die Praxis zu übersetzen, müssen Redaktionen nun auch Public-Interest-Angebote bewerten: Welche sind wirklich nachhaltig? Welche transparent? Welchen Governance-Modellen kann man trauen? Ohne klare institutionelle Strukturen droht aus der gemeinwohlorientierten Alternative sonst eine zusätzliche Unsicherheitsquelle zu werden.

Ob dieser konzeptionelle Diskurs um Public-Interest-KI in der deutschen Medienpraxis überhaupt ankommt, zeigt sich exemplarisch im folgenden Abschnitt. Dort wird deutlich: Während auf europäischer Ebene über gemeinwohlorientierte Infrastrukturen debattiert

wird, kämpfen deutsche Redaktionen mit der konkreten Umsetzung des AI Acts – ohne dass Public-Interest-Alternativen – bisher – eine wahrnehmbare Rolle spielen. Die Lücke zwischen gemeinwohlorientiertem Diskurs und redaktionellem Alltag ist eklatant.

5.4 Deutschland als Anwendungsfall: Umsetzung des AI Acts und praktische Unsicherheiten in Redaktionen

Während auf europäischer Ebene über Public-Interest-KI als gemeinwohlorientierte Alternative debattiert wird, folgt die deutsche Umsetzung des AI Acts einer anderen Logik: Die Wirkung der Regulierung konkretisiert sich bislang weniger durch konkrete Eingriffe in die redaktionelle Praxis als durch institutionelle Zuständigkeiten, Auslegungen und Stellungnahmen zentraler Akteur:innen. Mit der Bundesnetzagentur als vorgesehener Durchsetzungsbehörde wird KI-Regulierung primär als technisch-administrative Aufgabe verankert – Public-Interest-Alternativen spielen dabei keine Rolle, obwohl auch mehr als eineinhalb Jahre nach Inkrafttreten des EU AI Acts im August 2024 noch kein nationaler Durchführungsgesetzesentwurf vorliegt.¹²¹

Diese Verortung prägt auch den öffentlichen Diskurs: Im Vordergrund stehen Fragen der Konformität, Marktaufsicht und Systemrisiken. Die redaktionelle Praxis erscheint in dieser Perspektive allenfalls mittelbar betroffen und wird nicht als eigenständiger Regelungsadressat behandelt. Dies wird in den Stellungnahmen journalistischer Verbände besonders deutlich: Der Bundesverband Digitalpublisher und Zeitungsverleger (BDZV) und der Medienverband der freien Presse (MVFP) verweisen in ihren Positionen und Branchenerhebungen darauf, dass KI in deutschen Redaktionen breit eingesetzt wird, jedoch häufig ohne ausreichende rechtliche und organisationale Absicherung.

In den Positionen wird dabei wiederholt darauf hingewiesen, dass dieser Einsatz häufig schneller voranschreitet als die interne Klärung von Zuständigkeiten, Haftungsfragen und Prüfprozessen mithält: Der Einsatz von KI wird als notwendig beschrieben, um ökonomischer Schieflage, Effizienzanforderungen und veränderten Publikumsgewohnheiten zu begegnen.¹²² Der AI Act wird aus Sicht der Verlage insoweit zwar als notwendiger Ordnungsrahmen anerkannt, bietet für redaktionelle Anwendungen aber kaum konkrete Entscheidungshilfen. Fragen der Haftung, Transparenzpflichten oder Abgrenzung zulässiger Anwendungen bleiben aus Sicht der Verlage offen und müssen jeweils aufwändig redaktionsintern aufwändig geklärt werden, wodurch Regulierung hier tendenziell als zusätzlicher Stressfaktor erscheint.¹²³

Aus arbeits- und tarifpolitischer Perspektive formuliert dagegen ver.di eine komplementäre, aber anders akzentuierte Diagnose: In Stellungnahmen und Verhandlungen wird der

Einsatz von KI als Thema wachsender Relevanz benannt, gerade weil externe Regulierung bislang keine ausreichende Klarheit für den redaktionellen Betrieb schafft. Die Gewerkschaft verweist auf die Gefahr, dass Verantwortung und Kontrolle in KI-gestützten Arbeitsumgebungen unklar bleiben und Belastungen auf einzelne Beschäftigte – und mitunter auch Freie – verlagert werden.¹²⁴ Genauer betrifft dies Fragen der inhaltlichen Verantwortung, der Prüfpflichten sowie des Umgangs mit Fehlern in KI-gestützten Arbeitsabläufen vor allem unter Zeitdruck und Ressourcenknappheit. Der Ruf nach verbindlichen betrieblichen Regelungen mutet damit wie eine Reaktion auf eine Regulierungsmaßnahme an, die abstrakte Leitplanken setzt, ohne die konkrete Arbeitssituation in Redaktionen ausreichend zu kennen.

Da der AI Act einen rechtlichen Rahmen etabliert, zentrale Fragen der praktischen Umsetzung jedoch offenlässt, verlagert sich ein erheblicher Teil der Klärungsarbeit somit auf Verbände, Medienunternehmen und einzelne Redaktionen. Public-Interest-KI spielt in diesen Stellungnahmen keine Rolle. Weder BDZV noch ver.di fordern alternative Infrastrukturen – ihre Forderungen konzentrieren sich auf Compliance, Haftung und betriebliche Regelungen innerhalb proprietärer Systeme. Resilienz entsteht hier also nicht durch regulatorische Eindeutigkeit, sondern durch kollektive Übersetzungsarbeit zwischen geltendem Recht, betrieblichen Vorgaben und professionellem Anwendungsbezug.

Diese Übersetzungsarbeit ist mühsam und bislang sehr ungleich verteilt: Während größere Medienhäuser in der Lage sind, regulatorische Anforderungen juristisch zu bewerten und interne Prozesse anzupassen (etwa durch große Rechtsabteilungen, Compliance-Strukturen oder spezialisierte KI-Teams), stoßen kleinere Redaktionen schneller an Grenzen. Hier zeigt sich, was in Kapitel 5.3 bereits als strukturelles Defizit beschrieben wurde: Ohne gemeinwohlorientierte Infrastrukturen, die allen Redaktionen unabhängig von ihrer Größe zur Verfügung stehen, verschärft die AI-Act-Umsetzung bestehende Ungleichheiten. Eine Public-Interest-KI könnte diese Lücke schließen – wird aber in der deutschen Medienpolitik nicht ausreichend als Handlungsoption diskutiert.

Die Verbandspositionen machen zudem deutlich, dass KI-Regulierung nicht neutral wirkt, sondern bestehende strukturelle Unterschiede im Mediensystem noch weiter verstärken kann. Was insgesamt fehlt: eine Medienpolitik, die Regulierung und gemeinwohlorientierte Infrastruktur zusammendenkt. Die zuvor beschriebene Asymmetrie zwischen verbindlicher Regulierung und voluntaristischer Public-Interest-KI wird gerade in Deutschland deutlich. Während der AI Act umgesetzt wird, bleiben alternative Infrastrukturen ein Randthema – obwohl sie gerade für kleinere Redaktionen entscheidend wären. An dieser Schnittstelle entstehen Zielkonflikte zwischen Standardisierung, Vielfalt und Innovationsfähigkeit, die abschließend in den Blick genommen werden.

5.5 Konfliktlinien und Zielkonflikte:

Freiheit, Schutz und Innovationsfähigkeit

Die bisherigen Abschnitte haben deutlich gemacht, dass Regulierung im Kontext von KI nicht als eindimensionale Steuerungsaufgabe verstanden werden darf. Sie bewegt sich im Journalismus vielmehr in einem Feld struktureller Zielkonflikte, die sich nicht auflösen lassen, sondern politisch und professionell ausgehalten werden müssen. Die regulatorischen Vorgaben auf EU-Ebene und ihre nationale Umsetzung markieren einen Versuch, diesen Unwuchten einen rechtlichen Rahmen zu geben. Gleichzeitig zeigen sie, dass Regulierung stets selektiv wirkt und mindestens vier neue Konfliktlinien hervorbringt.

1. Ein zentraler Zielkonflikt betrifft das **Verhältnis von Schutz und Freiheit**. Der AI Act ist *grundrechtlich* motiviert. Er zielt darauf ab, Risiken für individuelle Rechte, gesellschaftliche Teilhabe und demokratische Prozesse zu begrenzen. Für den Journalismus ist dieser Schutzanspruch grundsätzlich anschlussfähig, da journalistische Arbeit selbst auf der Wahrung von Grundrechten beruht. Zugleich entsteht ein Spannungsfeld dort, wo regulatorische Schutzlogiken in Bereiche hineinwirken, die traditionell durch *redaktionelle Autonomie* geprägt sind. So kann etwa die Pflicht zur Dokumentation von KI-Einsätzen in Konflikt mit investigativem Arbeiten geraten: Nutzt eine Redaktion ein KI-Tool zur Auswertung sensibler Leaks, kann eine lückenlose Protokollierung der Eingaben u.U. Rückschlüsse auf geschützte Quellen ermöglichen und damit den Quellenschutz unterlaufen. Ähnlich kollidieren transparente Prüf- und Genehmigungsverfahren mit der Logik schneller publizistischer Entscheidungen, etwa wenn in einer Eil- oder Krisensituation in Minuten über die Nutzung KI-vorausgewählter Inhalte entschieden werden muss. An solchen Punkten zeigt sich exemplarisch, dass rechtliche Schutzmechanismen auf formalisierte Abläufe zielen, während journalistische Praxis gerade in ihren demokratierelevanten Spitzenleistungen häufig auf informelle, schwer regulierbare Handlungsweisen angewiesen ist.
2. Ein weiterer Konflikt liegt im **Verhältnis von Rechtssicherheit und Innovationsfähigkeit**, gerade bei schnellen Innovationsdynamiken wie im Falle der KI-Entwicklung. Weil viele Vorgaben allerdings sehr abstrakt formuliert sind, wissen Redaktionen oft nicht genau, ob und wie bestimmte KI-Tools (etwa für Überschriften, Bildbearbeitung oder automatische Meldungen) darunterfallen. Ohne klare Expertise und Erfahrungen schätzen sie Risiken eher als zu hoch ein. Dadurch entsteht eine vorsichtige, teils gebremste Haltung gegenüber KI: Vor allem kleinere Redaktionen verzichten eher auf offizielle Experimente oder Pilotprojekte und greifen stattdessen auf inoffizielle,

individuelle Lösungen einzelner Mitarbeitender zurück. Nicht die Regulierung selbst hemmt daher Innovation, sondern die Unsicherheit darüber, wie sie konkret anzuwenden ist.

3. Eng damit verbunden ist der Konflikt zwischen **Standardisierung und Vielfalt**. Der AI Act arbeitet mit generalisierenden Kategorien und einheitlichen Anforderungen, um europaweit vergleichbare Standards zu setzen. Auch der Journalismus kennt gemeinsame professionelle Normen, etwa zu Sorgfalt und Transparenz, setzt sie aber in sehr unterschiedlichen Arbeitsabläufen, Organisationsformen und mit sehr unterschiedlichen Ressourcen um. Eine einheitliche regulatorische Logik trifft daher auf ungleiche Voraussetzungen. Große Medienhäuser können Vorgaben zu Dokumentation oder Risikobewertungen in vorhandene Compliance-Strukturen und technische Systeme einbauen. Eine kleine Lokalredaktion mit wenigen Vollzeitstellen muss dieselben Pflichten neben dem Tagesgeschäft erfüllen – etwa zusätzliche Protokolle zur KI-Nutzung führen und Schulungen organisieren, ohne Rechtsabteilung oder eigene IT. So wird Regulierung leicht zu einer überproportionalen Belastung für kleinere und mittelgroße Redaktionen und kann bestehende Ungleichheiten im Mediensystem unbeabsichtigt verstärken.
4. Ein weiterer Zielkonflikt betrifft die **Verlagerung von Verantwortung**. Regulierung setzt auf klare Zuweisung von Pflichten und Haftung, um Risiken beherrschbar zu machen. In KI-gestützten journalistischen Arbeitsabläufen bleibt Verantwortung jedoch häufig verteilt. Entscheidungen entstehen im Zusammenspiel von Redaktion, Technik, externen Dienstleister:innen und Plattformen (vgl. Abb. 14). Der Versuch, Verantwortung rechtlich eindeutig zu fixieren, steht dieser prozessualen Realität diametral entgegen. Für Redaktionen kann das bedeutet, dass formale Verantwortlichkeit und faktische Kontrolle oder publizistische Haftung auseinanderfallen können. Resilienz entsteht hier bestenfalls durch die Fähigkeit, klug mit diesen disparaten Verantwortlichkeiten umzugehen.

Diese Zielkonflikte markieren die Grenzen medienpolitischer Steuerung in einem Feld, das von vielerlei Dynamiken, Unsicherheiten und Machtasymmetrien geprägt ist. Hinzu kommt ein weiterer, struktureller Zielkonflikt: das Verhältnis von Regulierung und Infrastruktur. Kapitel 5.3 hat gezeigt, dass Public-Interest-KI als gemeinwohlorientierte Ermöglichungsstrategie theoretisch zur Resilienz beitragen könnte – praktisch aber an begrifflicher Diffusion, fehlender Governance und mangelnder Nachhaltigkeit scheitert. Das deutet darauf hin, dass die aktuelle Medienpolitik nicht in der Lage ist, bestehende Friktionen zu lösen, ohne in ihren Handlungen selbst neue Probleme zu erzeugen. Ihre Aufgabe

besteht darin, Rahmenbedingungen zu setzen, innerhalb derer journalistische Praxis umgesetzt werden kann, ohne sie in technokratische Kontrolllogiken zu pressen. Regulierung schafft also notwendige Voraussetzungen, liefert aber keine Antworten auf die Frage, wie Resilienz im redaktionellen Alltag konkretisiert wird.

KI-Resilienz im Journalismus als geteilte Gestaltungsaufgabe

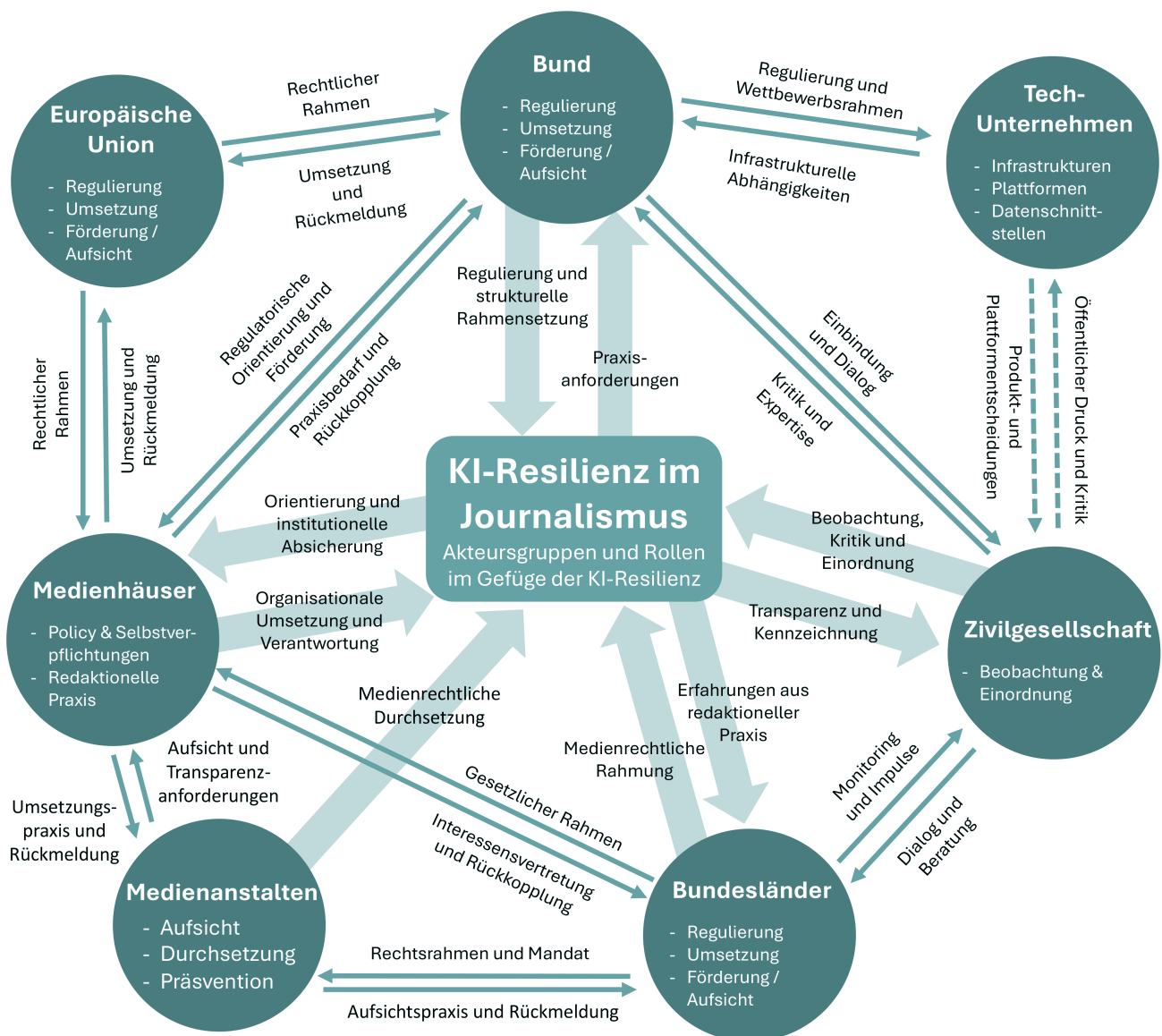


Abbildung 14: KI-Resilienz im Journalismus als geteilte Gestaltungsaufgabe – Wechselbeziehungen zwischen EU, Bund und Ländern (Regulierung und Förderung), Medienanstalten und Medienhäusern (organisationale Umsetzung), Tech-Unternehmen (Infrastruktur, Plattformen) und Zivilgesellschaft (Beobachtung, Kritik, Beratung) (eigene Darstellung)

5.6 Regulierung setzt Grenzen – Resilienz entsteht in der Praxis

Mit dem EU AI Act existiert erstmals ein verbindlicher europäischer Rechtsrahmen, der den Einsatz von KI grundrechtlich absichern und systemische Risiken begrenzen soll. Für journalistische Organisationen verändert dieser Rahmen die Bedingungen redaktioneller Praxis spürbar, auch wenn der Journalismus nicht als eigenständiger Regelungsadressat im Zentrum der Verordnung steht. Regulierung wirkt hier im Hintergrund, gleichwohl folgenreich: Sie definiert, was zulässig ist, welche Risiken zu vermeiden sind und welche Sorgfaltspflichten der europäische Gesetzgeber von Anbieter:innen und Anwender:innen von KI-Systemen – einschließlich journalistischer Organisationen – künftig erwartet. Sie schließt damit jedoch nicht die Lücke zwischen rechtlichem Rahmen und gelebter Praxis.

Diese Lücke füllt auch Public-Interest-KI nur bedingt aus. Zwar könnte eine gemeinwohlorientierte KI-Infrastruktur kleinere Redaktionen entlasten – doch solange sie begrifflich diffus und institutionell unverbindlich bleibt, verschärft die Asymmetrie zwischen verbindlicher Regulierung und voluntaristischer Förderlogik bestehende Ungleichheiten eher, als dass sie sie abbaut.

Die Grenze regulatorischer Steuerungsfähigkeit wird auch in aktuellen internationalen Analysen zur digitalen Öffentlichkeit sichtbar. Der UNESCO-Welttrendbericht von 2025 zur Freiheit der Meinungsäußerung und Medienentwicklung¹²⁵ zeichnet das Bild eines von Plattformlogiken, KI-gestützten Verbreitungsmechanismen und ökonomischen Konzentrationsprozessen geprägten Informationsökosystems, in dem Informationsintegrität, Vertrauen und demokratische Urteilsbildung unter Druck geraten – und sich nicht durch einzelne technologische Maßnahmen oder singuläre Akteure allein stabilisieren lassen.

Vor diesem Hintergrund wird deutlich, dass demokratische Öffentlichkeit bei weitem nicht allein durch rechtliche Normen oder technische Rahmensetzungen getragen werden kann. Sie bleibt stets auf vermittelnde Institutionen angewiesen – allen voran journalistische Medien, aber auch auf deren Selbstkontrollorgane, unabhängige Fact-Checking-Initiativen und zivilgesellschaftliche Organisationen, die Aufklärung betreiben, Konflikte sichtbar machen und Orientierung herstellen.¹²⁶

Damit wird deutlich, wie sich Regulierung und Resilienz zueinander verhalten: Der AI Act und begleitende nationale Umsetzungsprozesse definieren rechtliche Mindeststandards, sie markieren Grenzen und Verantwortlichkeiten. Im Sinne des hier entwickelten Resilienzbegriffs adressieren sie vor allem die *Dimension des Widerstands*: Die Prozesse zur Umsetzung des AI Acts sollen verhindern, dass KI journalistische Verantwortung, Transparenz oder Grundrechte unterminiert.

Gerade darin liegt sowohl ihre strukturelle Grenze als auch ein produktiver Bezugspunkt. Resilienz im Journalismus entsteht nicht im Gesetzestext, sondern in der Übertragung rechtlicher und ethischer Vorgaben in die gelebte Praxis. Sie entscheidet sich dort, *wo Redaktionen klären, wer unter KI-Bedingungen entscheidet, wie Verantwortung zugerechnet wird, wie mit Fehlern umgegangen wird und wie Unsicherheit strukturell bearbeitbar bleibt.*

Als Zwischenfazit: Regulierung schafft Verbindlichkeit, aber nicht zwingend Handlungskompetenz. In dieser Grauzone entscheidet sich, ob KI-Einsatz im Journalismus verantwortbar bleibt oder zur Quelle systemischer Fragilität wird – nicht durch den spektakulären Einzelfall, sondern durch die schleichende Verschiebung redaktioneller Entscheidungslogiken. Genau hier setzt Kapitel 6 an: Es fragt nicht mehr nach dem *Ob* oder *Warum* der Regulierung, sondern nach dem *Wie* ihrer Umsetzung – und will den Journalist:innen mit einem *praxisorientierten KI-Resilienz-Framework* konkrete Ansätze für eine souveräne Redaktionskultur an die Hand geben – mit dem Ziel, journalistische Urteilkraft, Verantwortung und Beziehung zum Publikum in belastbaren Entscheidungsarchitekturen systematisch zu verankern. .

.

6. KI-Framework für Redaktionen – Kein Toolkit für die Medienpraxis!

Worauf dieses Kapitel Antworten gibt:

1. Wie sich KI konkret, verantwortungsvoll und resilient in redaktionelle Arbeitsprozesse integrieren lässt.
2. Wie Redaktionen unter KI-Bedingungen entscheidungs- und handlungsfähig bleiben – jenseits von Tool-Fetischismus und Abwehrlogiken.
3. Welche internationalen und deutschen Praxisbeispiele gelungene – und gescheiterte – Formen redaktioneller KI-Governance zeigen.
4. Was der Realitätscheck dieser Beispiele mit Blick auf Verantwortungszuweisung, Transparenz und das „KI-Resilienz-Framework“ zeigt.
5. Wie KI-Nutzung Beziehungen verändert: innerhalb der Redaktion, zwischen Redaktion und Management – und zwischen Journalismus und Publikum.

6.1 Warum Frameworks? Orientierungswissen statt Tool-Partys

Die Debatte über KI im Journalismus wird bislang vor allem als Tool-Debatte geführt. Welche Anwendungen sind erlaubt? Welche Instrumente sind effizient? Welche sparen Zeit, Kosten, Personalstellen? Diese Fragen sind verständlich – und legitim. Aber sie greifen zu kurz. Denn sie verkennen, dass KI längst mehr ist als ein weiteres Tool im redaktionellen Werkzeugkasten. Sie ist zur *strukturellen Bedingung journalistischer Arbeit* geworden – oder wird es zumindest immer mehr.

Und zwar nicht nur im Offensichtlichen (z.B. Text- und Bildgenerierung, KI-Avatare), sondern auch im Verborgenen: bei Priorisierung, Zusammenfassung, Übersetzung, Moderation, SEO-Optimierung, Personalisierung, Recherchen und Workflow-Automatisierung. Damit wandert KI aus der Nische einiger ‚KI-Nerds‘ in die *Schaltzentralen redaktioneller Entscheidungen*. Wer hier weiterhin nur nach ‚erlaubt vs. verboten‘ fragt, übersieht die eigentliche (stille) Verschiebung: von der *Bedienungskompetenz zur Organisationsverantwortung*.

Genau deshalb braucht es, so glauben wir, einen *Perspektivwechsel* auf das Thema KI: Nicht das nächste Toolkit entscheidet darüber, ob KI den Journalismus stärkt oder aushöhlt – sondern die Frage, wie Redaktionen unter Zeit-, Verantwortungs- und Technologie- druck noch verlässlich urteilen, prüfen und begründen können. Im Kern geht es um die Stabilisierung redaktioneller Entscheidungsqualität unter neuen Bedingungen: mehr

Output-Dynamik, mehr synthetische Inhalte, mehr Unsicherheit über Herkunft und Validität von Inhalten.

Hinzu kommt eine strukturelle Bedingung, die in Kapitel 5 analysiert wurde: Redaktionen operieren in einem Umfeld fragmentierter Infrastrukturen. Während proprietäre Systeme dominieren, fehlen offene KI-Modelle, die dem Gemeinwohl verschrieben werden. Public-Interest-KI könnte diese Lücke vermutlich irgendwann schließen – bis dahin scheint es aber noch ein weiter Weg zu sein. Frameworks müssen daher bis auf Weiteres unter Bedingungen struktureller Abhängigkeit funktionieren können. Denn sie werden diese Abhängigkeit nicht auflösen, können aber Redaktionen befähigen, innerhalb dieser Grenzen verantwortungsvoller zu handeln.

Hier setzt unser Gedanke eines *KI-Resilienz-Frameworks* an. Frameworks sind keine KI-Toolkits, keine Bedienungsanleitungen und keine Best-Practice-Sammlungen im eigentlichen Sinn. Sie sind auch keine ‚Innovationsfolie‘ für Präsentationen, sondern ein konkretes Arbeitsinstrument, das Streitfragen im Alltag belastbar und umsetzbar macht. Und sie beantworten nicht die naheliegende Frage *Welches Tool sollen wir nutzen?*, sondern adressiert die wesentlich anspruchsvolleren Herausforderungen:

- *Wie wollen wir entscheiden, wenn sich technische Gelegenheiten und eine Nachrichtenlage unter Berichterstattungsdruck überlagern?*
- *Welche Mindeststandards gelten in solchen Situationen – auch im Ausnahme- oder Krisenzustand?*
- *Und wie sichern wir dabei journalistische Urteilskraft, Verantwortlichkeiten und Rechenschaft – gerade dann, wenn KI-Ergebnisse plausibel wirken?*

Diese Entscheidungsarchitekturen müssen zweierlei leisten: Sie müssen robust (resilient) genug sein, um auch beim nächsten Tool-Shift tragfähig zu sein, und konkret genug, um im Tagesgeschäft nicht als bloße ‚Haltungsangelegenheit‘ zu verpuffen. Der Prüfstein ist nicht das gut gemeinte Strategiepapier, sondern sozusagen die Nachtschicht unter realen Bedingungen, also: Was gilt, wenn es schnell gehen muss, wenn niemand ‚zuständig‘ sein will oder wenn das KI-Ergebnis verdächtig gut klingt?

Der Reiz vieler KI-Tools liegt dabei in ihrer scheinbaren Einfachheit: ein Prompt, ein Klick, ein plausibel klingendes Ergebnis. Diese Niedrigschwelligkeit verführt dazu, Entscheidungen zu individualisieren und zu entpolitisieren: Jede und jeder probiert ein wenig herum, passt sich an, findet eigene Workarounds. Doch genau darin liegt das Risiko: Die Redaktion verliert das, was sie – zumindest noch – selbst gegenüber Plattformlogiken auszeichnet: gemeinsame Standards, kollektive Verantwortung und überprüfbare Entscheidungen. Als besonders ambivalent sind die positiven und negativen psychologischen Effekte

einzustufen, die zunehmend die Bequemlichkeit in den Redaktionen fördern und deren Urteilkraft allmählich schwinden lassen.¹²⁷

Im Ergebnis entsteht ein *semiprofessioneller Schattenbetrieb*: viele kleinere, intransparente Praktiken – aber kein gemeinsames Regelwerk dafür, was ‚sicher‘, ‚vertretbar‘ und ‚journalistisch sauber‘ ist. Und dieser Schattenbetrieb ist nicht nur ein Compliance-Thema, sondern entlarvt auch ein Qualitätsproblem: Er erzeugt inkonsistente Tonalitäten, unterschiedliche Prüfintensitäten, unklare Befehlsketten und im Zweifel eine neue, zuweilen opake Fehlerklasse, die niemand mehr nachvollziehen kann.

Was fehlt, ist ein gemeinsamer praktischer Bezugsrahmen: ein KI-Resilienz-Framework, das Entscheidungen nach innen wie nach außen anschlussfähig macht. Anschlussfähig meint mehr als interne Abstimmung: Entscheidungen müssen erklärbar, verteidigungsfähig und wiederholbar sein – gegenüber Kolleg:innen, der redaktionellen Führungsebene, dem Publikum und im Zweifel auch gegenüber externen Prüfungen (Redaktionsbeirat, Ethikkommission, Medienkritik). Abbildung 15 markiert diesen Übergang von individueller und organisationaler Resilienz zu digitaler Souveränität als institutionelle Leistung.

Ein KI-Resilienz-Framework zielt somit nicht auf Effizienzversprechen, sondern primär auf redaktionelle Verlässlichkeit unter volatiler werdenden Bedingungen. Es immunisiert die Redaktionen nicht gegenüber handwerklichen oder berufsethischen Fehlern im Umgang mit KI, kann sie aber weniger anfällig dafür machen.

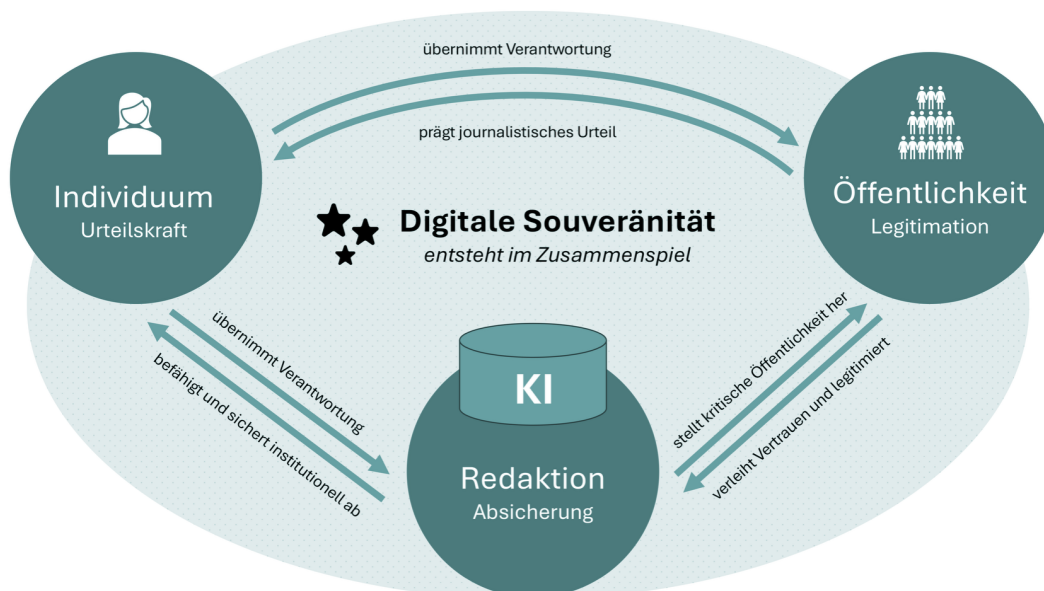


Abbildung 15: Digitale Souveränität im Zusammenspiel: Individuelle Urteilkraft, redaktionelle Absicherung und kritische Öffentlichkeit bilden gemeinsam den Rahmen, in dem der Einsatz von KI verantwortungsvoll, vertrauensvoll und journalistisch legitimiert funktionieren kann (eigene Darstellung)

Was ein KI-Resilienz-Framework leistet

Ein solches KI-Resilienz-Framework leistet folgendes: Es übersetzt abstrakte Prinzipien wie Transparenz, Verantwortung oder *Human-in-the-Loop* in strukturierte und gut nachvollziehbare Entscheidungslogiken und übersetzt diese in konkrete Handlungsketten: *Wer entscheidet was, zu welchem Zeitpunkt, mit welcher Eskalationslogik?* Es hilft Redaktionen, wiederkehrende Situationen einzuordnen: *Wann ist KI-Unterstützung überhaupt sinnvoll? Welches Problem kann ich identifizieren – und wie hilft mir die KI bei der Lösungsfindung? Wo sind im Umgang mit KI klare Grenzen nötig? Wer trägt welche Verantwortung bei welchem Schritt? Und wie lassen sich Effizienzgewinne realisieren, ohne journalistische Kernwerte zu beschädigen?*

Kurzum: Frameworks haben das Potenzial, aus KI-Nutzung eine konsistente organisierte Praxis zu machen – statt einer Ansammlung individueller Routinen. Entscheidend dabei ist, dass diese Praxis nicht auf Einzelfälle zielt, sondern auf wiederkehrende Muster. Frameworks schaffen Wiedererkennbarkeit im Redaktionsalltag, der aus unzähligen Einzelentscheidungen besteht. Sie entlasten damit nicht nur einzelne Akteur:innen wie etwa das redaktionelle Führungspersonal, sondern *die Redaktion als Ganzes* – ganz einfach, weil nicht jede Entscheidung neu ausgehandelt werden muss.

Wichtig ist dabei eine klare Abgrenzung: Dieses Whitepaper verfolgt bewusst keinen toolzentrierten Ansatz – nicht etwa aus Technikaversion, sondern aus professionell-empirischer Nüchternheit. Es geht also nicht darum, konkrete Anwendungen zu empfehlen oder zu verteufeln – oder technische Normen für immer festzuschreiben. Solche Empfehlungen sind ohnehin kurzlebig und folgen Marktzyklen, die sich schneller ändern, als redaktionelle Strukturen reagieren können. Sie werden auch der journalistischen Arbeit nicht gerecht. Denn wer KI ausschließlich über Tools definiert, hinkt zwangsläufig immer hinterher.

Unser Framework-Ansatz setzt erheblich tiefer an. Er fragt nicht nach dem *Was*, sondern nach dem *Wie* und *Wer*: *Wie kommen Entscheidungen zustande? Wer verantwortet sie? Und wie bleiben sie überprüfbar, auch wenn sich technische Voraussetzungen verändern?* Diese Fragen sind nicht technischer Natur, sondern organisatorischer – und damit genuin redaktionell. Unser Resilienz-Ansatz zielt also auf etwas Dauerhafteres: *auf organisationale Urteilskraft, redaktionelle Verantwortung und eine belastbare Fehlerkultur – nicht als Idealzustand, sondern als kontinuierliche Medienpraxis unter den aktuellen dynamischen Veränderungen.*

Strategische Entlastung statt technischer Überforderung

Genau hier zeigt sich der Unterschied zwischen kurzfristiger Anpassung und nachhaltiger Resilienz: Während Tool-Wissen schnell altert, bleibt die von uns gemeinte *Antizipations- und Entscheidungskompetenz auch längerfristig* anschlussfähig. Das KI-Resilienz-Framework schafft eine gewollte normative Distanz zur Technik, ohne Innovationsfähigkeit zu blockieren. Sie erkennen außerdem an, dass KI weder neutral noch autonom ist, sondern immer in soziale, ökonomische und institutionelle Kontexte eingebettet bleibt.

Frameworks verschieben den Fokus von technischer Sicherheit zu institutioneller Verlässlichkeit und helfen Redaktionen, technologische Veränderungen nicht jedes Mal als Ausnahmezustand zu erleben, sondern als wiederkehrende Herausforderung mit erprobten Entscheidungslogiken. Genau deshalb kann ihr Einsatz auch nicht rein operativ geregelt werden, sondern verlangt nach gezielter redaktioneller Steuerung (Governance). Für Redaktionen bedeutet das: KI-Nutzung wird zur strategisch-operativen Größe. Sie betrifft *Führungskultur* ebenso wie *Arbeitsprozesse*, die *journalistische Aus- und Weiterbildung* ebenso das *Verhältnis zum Publikum*.

Jede Entscheidung über KI ist damit auch eine Entscheidung darüber, wie Verantwortung umverteilt, Konflikte definiert und Öffentlichkeit hergestellt wird. Wer Frameworks etablieren möchte, muss zunächst eine Aussage darüber treffen, wie sich Journalismus per se positioniert – als bloßer Produktionsbetrieb oder als verantwortliche Institution im öffentlichen Raum. Frameworks sind daher nicht neutral, sondern normativ wirksam: Sie codieren ein bestimmtes *Verständnis von Professionalität*.

Ein weiterer Gedanke: Gerade unter Bedingungen knapper Ressourcen entfalten Frameworks ihre besondere Stärke. Sie schaffen Orientierung, ohne starre Regeln vorzugeben. Und sie verhindern, dass Redaktionen bei jeder neuen technologischen Welle in den reaktiven Modus der Überforderung (oder des Hypes) geraten. In diesem Sinne sind sie auch ein Instrument strategischer *Entlastung*, weil sie es ermöglichen, Entscheidungen zu bündeln, Prioritäten zu klären und begrenzte Ressourcen dort einzusetzen, wo sie journalistisch den größten Effekt erzielen. Effizienz entsteht hier aber nicht durch Beschleunigung und Masse um jeden Preis, sondern durch Klarheit darüber, was nicht ‚KI-fiziert‘ werden soll und darf.

Wichtig erscheint uns dabei: KI-Resilienz ist keine individuelle Kompetenz einzelner besonders technikaffiner Mitarbeitender, sondern eine *organisationale Fähigkeit, die ganze Teams, Arbeitsgruppen und die Redaktion als Ganzes betrifft*. Sie zeigt sich gerade nicht nur im Ausnahmefall, sondern im Regelbetrieb – also auch dort, wo Routinen greifen und Aufmerksamkeit knapp ist: Wenn Redaktionen Entscheidungen gemeinsam treffen, Verantwortung klar zuordnen und ihre Praxis transparent reflektieren, entsteht auch

Widerstandskraft gegen externe Abhängigkeiten und die Erosion innerhalb der Redaktion. Frameworks sind aus unserer Sicht deshalb das ideale Instrument, mit dem diese Fähigkeit am ehesten operationalisiert wird, indem sie abstrakte Werte in belastbare Entscheidungsabläufe übersetzen.

Mit diesem Anspruch führen wir im folgenden Unterkapitel die Idee eines *KI-Resilienz-Framework für Redaktionen* ein. Es versteht sich als *flexible Ordnungs- und Orientierungsinstanz* für unterschiedliche redaktionelle Kontexte, von der Lokalredaktion bis zum datenjournalistischen Tech-Team. Sein Zweck ist damit nicht Vereinheitlichung, sondern – im Gegenteil – eine Anschlussfähigkeit zwischen den vielgestaltigen Verhältnissen von Technik, Verantwortung und journalistischer Praxis zueinander. KI ist – wie zuvor diskutiert – in erster Linie eine Infrastrukturfrage, aus unserer Sicht aber auch eine Haltungsfrage – und diese beiden verlangen nach Orientierung, Governance und Berufsethik. Genau hier setzt das *KI-Resilienz-Framework* als strukturierte, offene Entscheidungsarchitektur für Redaktionen an.

6.2 Das KI-Resilienz-Framework für Redaktionen

Die bisherigen Kapitel haben gezeigt: Die Debatte über KI im Journalismus oszilliert bislang zwischen zwei verkürzten Perspektiven. Entweder wird KI als technisches Werkzeug behandelt, das mit Regeln, Schulungen und Tool-Listen unter Kontrolle gebracht werden soll und Auswege für allerlei Zwänge und Probleme verspricht. Oder sie erscheint als übermächtige Struktur, der sich die redaktionelle Arbeit entziehen oder zwangsläufig unterordnen muss. Beide Sichtweisen greifen nach unserer Erfahrung zu kurz. Sie unterschätzen entweder die kulturelle und organisationale Tiefe des Wandels – oder sie konterkarieren die Gestaltungsverantwortung. In beiden Fällen bleibt Journalismus lediglich *reaktiv*.

Was bislang fehlt, ist ein Ordnungs- und Orientierungsrahmen, der den Umgang mit KI in der journalistischen Praxis in seiner tatsächlichen Komplexität ernst nimmt.

Genau hier setzt das Verständnis von KI als infrastruktureller Herausforderung an. Der Begriff des ‚Frameworks‘ ist dabei bewusst gewählt: Ein Framework ist kein starres Regelwerk, sondern eine *Denkanleitung und Handlungsarchitektur*. Sie hilft Redaktionen, Entscheidungen besser zu strukturieren, Friktionen sichtbar zu machen und Verantwortung systematisch zu verorten – gerade dort, wo der Einsatz von KI nicht eindeutig als *richtig* oder *falsch* einzustufen ist. Frameworks dienen damit nicht der Kontrolle von oben, sondern der *professionellen Selbstwirksamkeit im redaktionellen Alltag*. Sie beantworten

nicht abschließend, *was zu tun ist*, sondern strukturieren für Redaktionen entscheidende Fragen vor:

- *Wo stehen wir?*
- *Nach welchen Kriterien entscheiden wir?*
- *Und wo entstehen neue Abhängigkeiten, ohne dass wir sie bewusst reflektieren?*

Das vorgestellte KI-Resilienz-Framework verfolgt damit ein klares Ziel: Redaktionen sollen in die Lage versetzt werden, KI nicht nur effizient, sondern verantwortungsvoll, transparent und beziehungsorientiert einzusetzen. Es geht ausdrücklich nicht um technologische Prozessoptimierung, sondern um *organisationale Handlungsfähigkeit*. KI-Resilienz entsteht – siehe Kapitel 3 – wiederum dort, wo Verantwortung nicht an Systeme delegiert wird, Transparenz nicht zur Pflichtübung verkommt und menschliche Beziehungspflege als journalistische Kernaufgabe verstanden wird.

Um diese Anforderungen in den redaktionellen Alltag zu übersetzen, gliedert sich das Framework in *sieben zentrale Handlungsfelder*. Diese Felder sind nicht additiv, sondern systemisch angelegt. Sie bündeln jene Bereiche, in denen sich die Verantwortlichkeiten, aber auch die Verwundbarkeiten des Journalismus im KI-Zeitalter besonders deutlich zeigen: im Verhältnis zum Publikum, in der Zuordnung von Haftung, im Schutz von Quellen, im Umgang mit Wahrheit und Fälschungen, in der Kompetenzverteilung innerhalb von Redaktionen und in der demokratischen Aufgabe des Journalismus insgesamt. Jedes Handlungsfeld adressiert einen eigenen Risikoraum – gemeinsam bilden sie jedoch *ein zusammenhängendes System mit ineinandergreifenden Handlungsfeldern*:

Transparenz ohne Verantwortung bleibt folgenlos. *Human-in-the-Loop* wird ohne Kompetenzaufbau zur bloßen Formalie. Technische Sicherheit ohne Beziehung untergräbt Vertrauen. Genau deshalb ist dieses Framework keine Checkliste im Sinne *technischer Compliance*. Es geht ausdrücklich nicht darum, Maßnahmen abzuhaken, sondern Zusammenhänge zu verstehen.

Unser Abgleich mit bestehenden KI-Leitlinien führender Medienhäuser zeigt (vgl. Kap. 4): Die hier vorgeschlagenen Handlungsfelder sind keine theoretische Neuerfindung, sondern eine *systematische Verdichtung realer Praxis*. Während viele Redaktionen einzelne Aspekte – etwa Transparenz, *Human-in-the-Loop* oder Schulungen/ Weiterbildungen – bereits aktiv adressieren, fehlt bislang ein integrierender Rahmen, der diese Maßnahmen miteinander verbindet.

Das KI-Resilienz-Framework setzt hier an: Die folgende Tabelle 1 überführt diesen systemischen Anspruch in eine praxisnahe Struktur. Entscheidend ist nicht, ob Redaktionen exakt diese Praktiken übernehmen, sondern wie sie diese in ihre jeweiligen Kontexte,

Kulturen und Entscheidungsabläufe integrieren. Ziel ist es, fragmentierte Leitlinien zu einer konsistenten redaktionellen Strategie zusammenzuführen – jenseits von Toolkits, einzelnen Ethikpapieren oder punktuellen Trainings. Vor diesem Hintergrund interessiert die Tabelle als zentrales Arbeitsinstrument: Sie markiert jene Stellen im redaktionellen Alltag, an denen sich entscheiden lässt, ob KI wirklich als Entlastung wirkt – oder als wiederkehrendes Reputationsrisiko. Auffällig ist, dass kaum eines der bislang veröffentlichten KI-Regelwerke großer Medienhäuser diese Entscheidungspunkte explizit miteinander verschränkt.

Das hier vorgeschlagene Framework schließt daher eine Lücke, indem es disparate Regeln und Maßnahmen in ein *kohärentes Orientierungsmodell* überführt. KI-Nutzung erscheint in dieser Perspektive nicht länger als Abfolge technischer Einzelfragen, sondern als *dynamische organisationale Gestaltungsaufgabe*.

Jede Entscheidung über Transparenz, Kontrolle, Datennutzung oder Automatisierung wirkt dabei auf mehreren Ebenen gleichzeitig: interne Arbeitskulturen, professionelle Selbstverständnisse und die Beziehung zum Publikum. Das Framework macht diese Wechselwirkungen nicht nur sichtbar, es schafft auch erst die Voraussetzung für eine reflektierte, souveräne und lernfähige KI-Praxis im Journalismus.

Die Tabelle erhebt dabei keinen Anspruch auf Vollständigkeit oder universelle Gültigkeit. Sie ist so gesehen keine Blaupause, sondern eine Denk- und Orientierungsstruktur, die größtenteils universellen Charakter hat, aber auch auf Bedürfnisse einzelner Medienhäuser oder Redaktionen zugeschnitten und individualisiert werden kann. Entscheidend ist es nicht einzelne Punkte zu befolgen, sondern sich bewusst mit den dahinterliegenden übergreifenden Fragen auseinanderzusetzen. Die entsprechenden Maßnahmen sind insoweit exemplarisch gewählt und sollen Redaktionen dabei unterstützen, eigene Lösungen zu entwickeln, Prioritäten zu setzen und implizite Entscheidungen explizit nach außen zu tragen.

Die Tabelle verbindet vier Dimensionen, die im Redaktionsalltag häufig getrennt behandelt werden: *Zielsetzung*, *Entscheidungslogik*, *empfohlene praktische Maßnahmen* und *Resilienz-Effekte*. Sie macht sichtbar, wie organisationale Resilienz unter KI-Bedingungen nicht via Einzelmaßnahmen, sondern durch deren systemisches Zusammenspiel gedeiht, und markiert damit den Übergang von der Analyse zur Gestaltung: von der Beschreibung bekannter Probleme hin zu einer strukturierenden Praxis, ohne deren systemische Komplexität zu verleugnen.

HANDLUNGSFELD	ZIEL	EMPFOHLENE MASS- NAHMEN	RESILIENZ-EFFEKT
1. TRANSPARENZ & KENNZEICHNUNG	Glaubwürdigkeit wahren, Publikum einbeziehen	Einheitliche Labels für KI-Inhalte; Disclosure-Boxen in Artikeln oder Beiträgen; jährliche Transparenzberichte	Stabilisierter Erwartungshorizont: Publikum kann KI-Einsatz einordnen, ohne journalistische Qualität grundsätzlich infrage zu stellen
2. HUMAN-IN-THE-LOOP	Fehler vermeiden, Verantwortung sichern	Vier-Augen-Prinzip; Redaktionsstatuten zu Review-Pfaden; klare Rollenklärung („KI-Editor:in“)	Verantwortung bleibt personalisiert und überprüfbar – auch bei teilautomatisierten Prozessen, stärkt journalistische Sorgfaltspflicht
3. QUELLEN- & DATENVERTRAUEN	Schutz von Informant:innen und Datenintegrität	<i>On-Premise</i> -Modelle; strikte Prompt-Hygiene; abgeschottete, verschlüsselte Datenräume	Vermeidung struktureller Abhängigkeiten von externen KI-Infrastrukturen bei sensiblen Informationen, fördert digitale Souveränität
4. FORENSIK & VERIFIKATION	Desinformation erkennen, Glaubwürdigkeit absichern	OSINT-Toolkits; Deepfake-Detection; Redaktions-„Red Teams“ für Simulationen	Institutionalisierte Prüfkompentenz ersetzt situative Plausibilitätsurteile, stärkt Abwehrfähigkeit gegen Manipulation und Fake-Kampagnen
5. ETHIK & ACCOUNTABILITY	Klare Verantwortung sichtbar machen	Interne Ethik-Boards; Fehler- und Korrekturmechanismen; transparente und offene Fehlerkommunikation	Förderung von Fehlerkultur, Verhinderung von Verantwortungsdiffusion zwischen Mensch, Organisation und System
6. KOMPETENZAUFBAU & TRAINING	Redaktionen kontinuierlich befähigen	Pflichtmodule zu KI-Resilienz und EU AI Act; interne KI-Wikis; Learning Sprints und KI-Resilience-Labs	Verankerung von Lernfähigkeit, verhindert Kompetenzgefälle innerhalb der Redaktion: Reduktion der Abhängigkeit von Einzelpersonen und externem Expertenwissen
7. PUBLIKUMSDIALOG & BEZIEHUNGSARBEIT	Beziehung zwischen Redaktion und Publikum stärken	Partizipative Formate (z. B. Ask-Me-Anything-Sessions zu KI-Einsatz); Feedback-Kanäle; Community-orientierte KI-Projekte	Fördert Nähe, Transparenz und langfristige Bindung an Medienmarken; Stabilisierung journalistischer Legitimität in einer zunehmend synthetischen Öffentlichkeit

Tabelle 1: KI-Resilienz-Framework für Redaktionen: Sieben Handlungsfelder (eigene Darstellung)

Die nachfolgend beschriebenen *sieben Handlungsfelder* markieren unterschiedliche Entscheidungsebenen, auf denen sich im KI-Zeitalter journalistische Verantwortung konkretisiert – auf denen jedoch zugleich neue Vulnerabilitäten aufscheinen. Sie adressieren technische Risiken, aber auch organisationale, kulturelle und gesellschaftliche Leitplanken. Sie lassen sich dabei jeweils als spezifische *Lesart von KI-Resilienz* verstehen (vgl. auch Abb. 16):

1. **Transparenz & Kennzeichnung:** Erwartungssicherheit und Vertrauensarbeit gegenüber dem Publikum.
2. **Human-in-the-Loop:** personalisierte Verantwortung unter automatisierten Bedingungen.
3. **Quellen- & Datenvertrauen:** journalistische Unabhängigkeit und digitale Souveränität.
4. **Forensik & Verifikation:** institutionalisierte Wahrheitsprüfung in einer synthetischen Öffentlichkeit.
5. **Ethik & Accountability:** Governance gegen Verantwortungsdiffusion.
6. **Kompetenzaufbau & Organisationslernen:** kollektive Urteilskraft statt individuelles Expertenwissen.
7. **Publikumsdialog & Beziehungsarbeit:** demokratische Legitimität über die Redaktion hinaus – hin zur demokratischen Funktion des Journalismus („Demokratie-Resilienz“).

Handlungsfeld 1: Transparenz & Kennzeichnung:

Warum Sichtbarkeit von KI-Einsatz zur neuen Vertrauensfrage wird

Transparenz war im Journalismus stets ein zentrales Vertrauensversprechen: Quellen offenlegen, Interessen benennen, Arbeitsweisen erklären, Fehler korrigieren. Mit dem Einsatz von KI verschiebt sich dieses Versprechen grundlegend. Denn KI wirkt häufig dort, wo sie nicht unmittelbar sichtbar ist – im Hintergrund von Recherche, Textproduktion, Bildbearbeitung, Datenanalyse oder Themenpriorisierung. Genau darin liegt das Problem: Was unsichtbar bleibt, entzieht sich der Beziehung zwischen Redaktion und Publikum.

Viele Redaktionen reagieren darauf mit Zurückhaltung. Sie fürchten, dass die Kennzeichnung von KI-Einsatz den Zweifel an journalistischer Qualität nähren oder Vertrauen untergraben könnte. Diese Sorge ist nachvollziehbar, greift jedoch zu kurz. Denn empirisch wie normativ gilt: *Das eigentliche Vertrauensrisiko entsteht nicht durch Offenlegung, sondern durch Intransparenz.* In einer Mediumgebung, die zunehmend von synthetischen Inhalten, automatisierten Prozessen und schwer unterscheidbaren Quellen geprägt ist, wird Nachvollziehbarkeit selbst zur zentralen Währung.

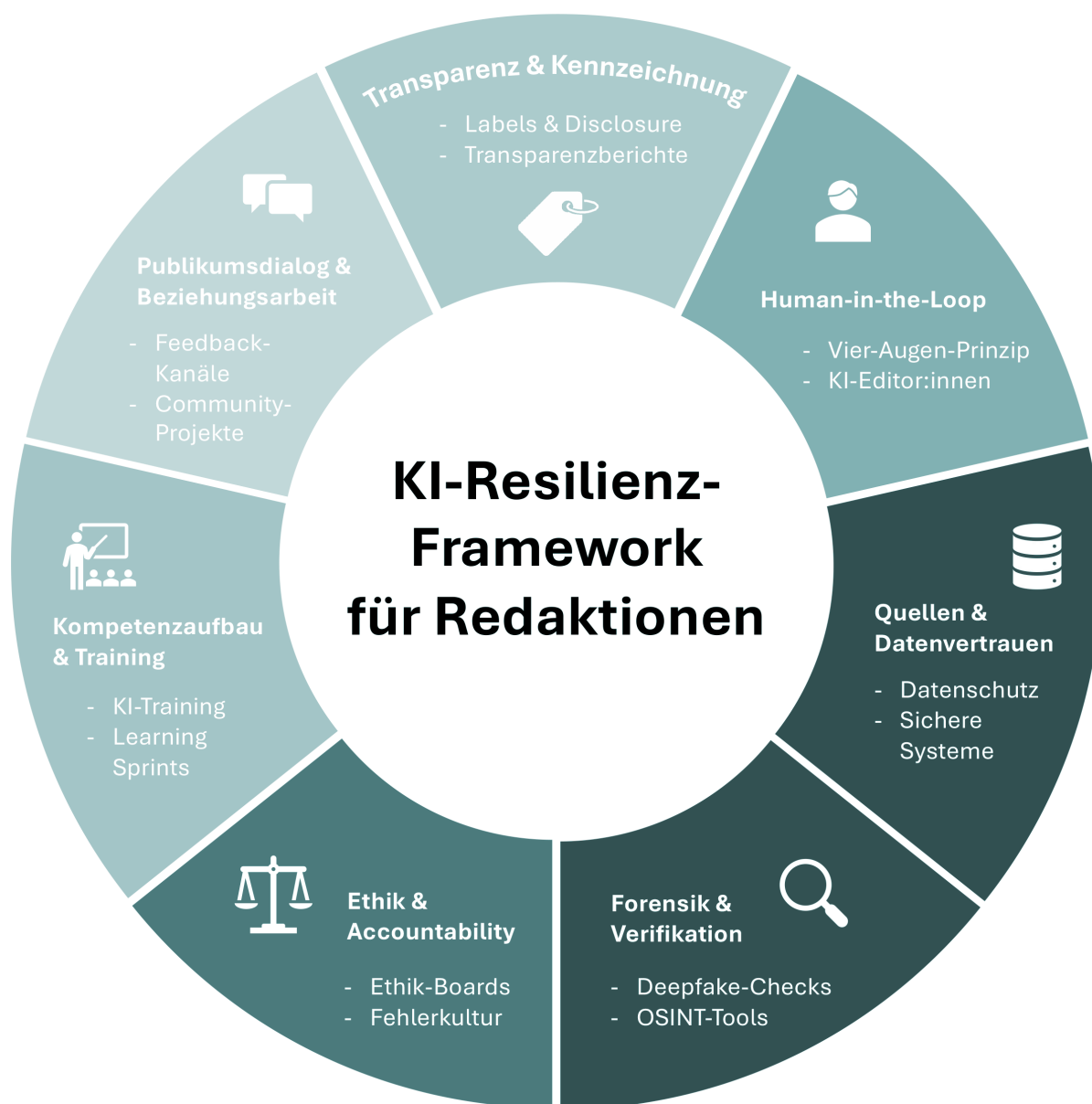


Abbildung 16: Das KI-Resilienz-Framework für Redaktionen: Sieben Handlungsfelder für organisationale Stabilität (eigene Darstellung)

Transparenz bedeutet dabei nicht, jede technische Detailfrage offenzulegen oder interne Workflows vollständig zu dokumentieren. Es geht nicht um Tool-Listen, Prompt-Archive oder Code-Offenlegungen im Artikel. Entscheidend ist vielmehr, dass das Publikum versteht, ob, wo und in welcher Rolle KI im journalistischen Prozess eingesetzt wurde – und

wo nicht. *Transparenz ist damit eine Frage der Einordnung, nicht der technischen Plausibilität.*

Einheitliche Kennzeichnungen für KI-unterstützte Inhalte, kurze Disclosure-Boxen oder öffentlich zugängliche Transparenzseiten schaffen genau diese Orientierung. Sie ermöglichen es dem Publikum, redaktionelle Entscheidungen besser einzuordnen, ohne Inhalte zu relativieren oder zu delegitimieren. Entscheidend ist dabei die Verlässlichkeit: Transparenz darf nicht situativ oder defensiv eingesetzt werden, sondern muss als konsistenter Standard erkennbar werden.

Wichtig ist es zudem, Transparenz nicht als reaktive Schadensbegrenzung zu begreifen, sondern als aktive Beziehungsarbeit. Wer offenlegt, signalisiert: Wir reflektieren unseren Technologieeinsatz, wir machen Entscheidungen sichtbar, und wir übernehmen Verantwortung – auch dort, wo Prozesse komplexer werden. Gerade in einer Phase, in der KI-Systeme zunehmend autoritativ wirken, wird diese Haltung zu einem zentralen Differenzierungsmerkmal journalistischer Glaubwürdigkeit.

Der Resilienz-Effekt von Transparenz liegt darin, Unsicherheit vorwegzunehmen, statt auf Vertrauenskrisen zu reagieren. Redaktionen, die ihren KI-Einsatz systematisch kenntlich machen, verschieben die Deutungshoheit zurück zu sich selbst. Sie definieren, was KI im eigenen journalistischen Kontext bedeutet, bevor diese Deutung von außen zugeschrieben wird. Transparenz verbindet technische Praxis mit normativer Verantwortung und öffnet den Raum für Dialog. Ohne Transparenz bleiben alle weiteren Handlungsfelder – Verantwortung, Verifikation, Ethik oder Kompetenzaufbau – für das Publikum nebulös. KI-Resilienz beginnt daher dort, wo Redaktionen bereit sind, ihre eigene Praxis – und auch ihre Fehler – nachvollziehbar zu machen.

Der hier vorgeschlagene Ansatz geht bewusst darüber hinaus. Regelmäßige, etwa jährliche Transparenzberichte zum KI-Einsatz machen sichtbar, wie sich Nutzung, Risiken und Routinen über die Zeit verändern. Transparenz wird so nicht reaktiv, sondern strukturell eingesetzt. Genau darin liegt ein zentrales Profilvermerkmal dieses Frameworks: *Es versteht Offenlegung nicht als Ausnahme, sondern als Teil redaktioneller Selbstverpflichtung.*

Handlungsfeld 2: Human-in-the-Loop und Verantwortungszuordnung:

Warum Verantwortung nicht automatisiert werden darf

Der Begriff ‚Human-in-the-Loop‘ wird in der KI-Debatte häufig technisch verkürzt verstanden: als letzter Kontrollblick eines Menschen auf ein maschinell erzeugtes Ergebnis. Für journalistische Praxis ist diese Terminologie jedoch missverständlich. Denn im

Journalismus geht es nicht primär um Kontrolle, sondern um Verantwortung – und um die Frage, wer entscheidet, wofür einsteht und wer im Zweifel rechenschaftspflichtig ist.

KI-Systeme sind darauf ausgelegt, Prozesse zu beschleunigen, Vorschläge zu generieren und Entscheidungen vorzustrukturieren. Genau darin liegt ihr praktischer Nutzen. Gleichzeitig entsteht ein strukturelles Risiko: Je plausibler, konsistenter und souveräner KI-Ergebnisse erscheinen, desto größer ist die Versuchung, menschliche Urteilkraft auf eine bloße Abnahmerolle zu reduzieren. Zustimmung wird zur Routine, Zweifel eher die Ausnahme. *Human-in-the-Loop* darf deshalb nicht als formaler Kontrollhaken am Ende des Produktionsprozesses ausgedeutet werden, sondern als *institutionalisierte Entscheidungspraxis*.

Für Redaktionen bedeutet das: Es braucht klar definierte Review-Pfade. Wer prüft KI-gestützte Inhalte? In welcher Phase des redaktionellen Workflows? Mit welcher Entscheidungsbefugnis? Und mit welcher Verantwortung bei Fehlern? Ohne solche Klärungen droht die Verantwortung zwischen Mensch und Maschine zu oszillieren – oder stillschweigend doch an das System delegiert zu werden.

Instrumente wie das Vier-Augen-Prinzip, verbindliche Freigaberegeln oder explizit definierte Rollenprofile – etwa KI-Editor:innen oder Review-Verantwortliche – sind in diesem Zusammenhang keine Bürokratie, sondern *organisationale Sicherheitsnetze*. Sie schaffen Klarheit darüber, dass KI zweifellos (gute) Vorschläge liefern kann, Entscheidungen, sie zu veröffentlichen, jedoch immer menschlich getroffen werden. Wegweisend ist nicht die formale Existenz solcher Rollen, sondern ihre tatsächliche Anwendung im redaktionellen Alltag.

Human-in-the-Loop ist also nicht nur eine Frage von Prozessen, sondern vor allem einer gefestigten Redaktionskultur. Das Prinzip funktioniert nur dort, wo Widerspruch ausdrücklich erlaubt – ja erwartet – ist. Journalist:innen müssen das Recht und die Pflicht haben, maschinelle Vorschläge zu hinterfragen, zu verwerfen oder zu überschreiben, ohne dies rechtfertigen zu müssen. Resiliente Redaktionen fördern genau diese Haltung: KI ist Assistenz, kein Menschenersatz.

Ein zentrales Risiko besteht jedoch darin, dass menschliche Kontrolle zwar formal vorgesehen ist, faktisch aber unter Zeitdruck, Ressourcenknappheit oder Routine ausgehöhlt werden kann. Wo Review-Schritte zur bloßen Formsache werden, verliert der *Human-in-the-Loop* seine Schutzfunktion. Deshalb ist es zielführend, Verantwortungsdiffusion nicht nur zu dokumentieren, sondern regelmäßig zu reflektieren: *Wo wird KI tatsächlich nachgeprüft? Wo wird sie nur noch abgenickt?*

Der zentrale Resilienz-Effekt dieses Handlungsfelds liegt deshalb in der Re-Personalisierung von Verantwortung: Operativ sichern klare *Human-in-the-Loop*-Strukturen journalistische Sorgfaltspflichten ab und reduzieren Fehlerrisiken. Kulturell stärken sie das professionelle Selbstverständnis und die Bereitschaft zum kritischen Urteil. Institutionell machen sie Verantwortung überprüfbar – nach innen wie nach außen. Sie signalisieren: Auch im KI-Zeitalter bleibt Journalismus eine menschliche Entscheidungspraxis, unterstützt durch Technologie, aber nicht von ihr maßgeblich gesteuert.

Innerhalb des Gesamtframeworks fungiert *Human-in-the-Loop* somit als Scharnier zwischen Technik und Ethik. Ohne klare Verantwortungszuordnung bleiben Transparenzmaßnahmen hohl und eine redaktionelle Fehlerkultur wirkungslos. Erst dort, wo menschliche Entscheidungsmacht explizit verankert ist, können Effizienzgewinne mit journalistischer Integrität verbunden werden. *Human-in-the-Loop* ist, im Ganzen gesehen, keine technische Option, sondern eine *demokratische Mindestbedingung des professionellen Journalismus*.

Die explizite Rollenklärung, wie sie das Framework mit der Idee eines ‚KI-Editors‘ oder vergleichbarer Verantwortlichkeiten vorschlägt, stellt daher einen qualitativen Fortschritt dar. Verantwortung wird hier eben nicht nur behauptet, sondern personell verortet – womit der *Human-in-the-Loop* von einer abstrakten Kontrollidee zu einer integrierten organisationalen Rolle wird.

Handlungsfeld 3: Quellen- & Datenvertrauen:

Warum digitale Souveränität bei sensiblen Informationen beginnt

Quellenschutz ist kein Nebenprodukt journalistischer Arbeit, sondern ihr Fundament. Anders gesagt: Ohne das Vertrauen von Informant:innen gäbe es keine investigativen Recherchen, ohne Datensicherheit keine glaubwürdige Redaktion. Mit dem Einsatz von KI geraten genau diese Grundlagen unter Druck – oft leise, indirekt und deshalb besonders gefährlich. Denn KI-Systeme verändern nicht nur, *was* verarbeitet wird, sondern *wo* und *wie* Informationen zirkulieren.

Viele heute verbreitete KI-Anwendungen arbeiten cloudbasiert, greifen auf externe Infrastrukturen zu und speichern Eingaben zumindest temporär. Für journalistische Arbeit ist das hochproblematisch. Recherchematerial, Interviewnotizen, vertrauliche Dokumente oder interne Analysen sollten nicht unkontrolliert in externe Systeme eingespeist werden – weder absichtlich noch aus Unachtsamkeit. Quellenvertrauen ist binär: Es besteht oder es wird – mitunter irreversibel – beschädigt.

KI-Resilienz beginnt deshalb bei der bewussten Gestaltung von Datenflüssen. Redaktionen müssen wissen, *welche Informationen in welche Systeme gelangen, wer Zugriff hat und was mit diesen Daten geschieht* – technisch wie organisatorisch. On-Premise-Modelle, abgeschottete Datenräume, klare Regeln zur Prompt-Hygiene und verbindliche Klassifikationen sensibler Inhalte sind keine optionalen Sicherheitsmaßnahmen, sondern Voraussetzungen journalistischer Integrität im KI-Zeitalter.

Dabei geht es nicht nur um IT-Sicherheit im engen Sinn, sondern um *digitale Souveränität*. Wer sich vollständig auf proprietäre Systeme verlässt, gibt Kontrolle ab – über Daten, über Arbeitsprozesse und über die eigene Unabhängigkeit. Die in Kapitel 5 diskutierten Ansätze für gemeinwohlorientierte Infrastrukturen – etwa GAIA-X oder Public-Interest-KI-Initiativen – könnten diese Abhängigkeit idealiter reduzieren. Doch die Umsetzungen zeigen: Solche Infrastrukturen sind fragmentiert, unterfinanziert und bieten bislang keine einsatzfertigen journalistischen Anwendungen.

Besonders heikel ist es allerdings dort, wo KI-Systeme implizit ‚lernen‘ oder journalistische Inputs für deren Trainings- oder Optimierungszwecke verwendet werden. Selbst wenn Anbieter versichern, keine Trainingsdaten zu speichern, bleibt für Redaktionen ein strukturelles Restrisiko bestehen. Quellen- und Datenvertrauen ist daher nicht nur eine juristische oder technische, sondern eine *politische Frage*. Sie berührt das Machtverhältnis zwischen Redaktion und Technologieanbietern ebenso wie das Verhältnis zwischen Redaktion und Quelle. Wer nicht garantieren kann, dass sensible Informationen geschützt bleiben, verliert seine Fähigkeit zur kritischen Berichterstattung – unabhängig von formalen Zusicherungen.

Ein oft unterschätzter Aspekt ist dabei die Rolle von Routinen. Viele Datenschutzverletzungen entstehen nicht durch böse Absicht, sondern durch Alltagspraxis: das schnelle Kopieren von Texten in ein KI-Tool, das beiläufige Zusammenfassen eines Interviews, das Hochladen eines Dokuments ‚nur zur Strukturierung‘. Ohne klare Regeln zur Prompt-Hygiene und entsprechende Schulungen entstehen so schleichende Risiken, die sich der individuellen Kontrolle entziehen.

Deshalb gehört Quellen- und Datenvertrauen untrennbar mit Kompetenzaufbau und Governance verknüpft. Technische Schutzmaßnahmen allein reichen nicht aus, wenn Redaktionen nicht wissen, wie KI-Systeme ihre Daten verarbeiten – und wo die Grenzen liegen. Umgekehrt bleiben Schulungen wirkungslos, wenn organisatorische Schutzräume fehlen. Resiliente Redaktionen denken Schutz und Kompetenz demnach zusammen. Der Resilienz-Effekt dieses Handlungsfelds liegt im Schutz des journalistischen Wesenskerns. Redaktionen, die Daten- und Quellensicherheit konsequent mitdenken, sichern ihre Fähigkeit zu investigativer, kritischer und unabhängiger Berichterstattung auch unter

KI-Bedingungen. Sie signalisieren nach innen wie nach außen: Effizienz darf niemals auf Kosten von Vertrauen gehen.

Im Gesamtframework fungiert Quellen- und Datenvertrauen als Schutzschild gegen strukturelle Abhängigkeiten. Ohne dieses Handlungsfeld verlieren Transparenz, *Human-in-the-Loop* und Ethik ihre Substanz. Denn wo Datenflüsse unkontrolliert sind, wird Verantwortung faktisch entkoppelt. KI-Resilienz bedeutet hier: Kontrolle über die sensiblen Informationen behalten – auch dann, wenn technologische Lösungen verlockend einfach erscheinen.

Die hier vorgeschlagenen Maßnahmen – von strikter Prompt-Hygiene über abgeschottete Datenräume bis hin zu *On-Premise-Modellen* – übersetzen klassische Prinzipien des Quellenschutzes erstmals konsequent in KI-Architekturen. Genau darin liegt ein zentrales Unterscheidungsmerkmal dieses Frameworks: Digitale Souveränität wird nicht postuliert, sondern praktisch gestaltet.

Handlungsfeld 4: Forensik & Verifikation:

Warum Wahrheitsprüfung zur Infrastrukturfrage wird

Forensik und Verifikation gehörten immer zum journalistischen Handwerkszeug. Quellen prüfen, Aussagen gegenrecherchieren, Material kontextualisieren – all das ist kein neues Geschäft. Doch mit dem Aufkommen generativer KI verschiebt sich die Bedeutung dieser Praktiken grundlegend. Was früher primär eine Frage von Sorgfalt, Erfahrung und Gegenchecks war, wird heute zu einer permanenten infrastrukturellen Herausforderung.

Denn KI produziert nicht nur falsche Informationen. Sie produziert plausible Fälschungen. Deepfakes, synthetische Stimmen, KI-generierte Bilder, Texte und Videos unterlaufen klassische Plausibilitätsmarker. Fehler sind nicht mehr grob oder leicht erkennbar, sondern sprachlich sauber, visuell überzeugend und kontextuell anschlussfähig. Genau darin liegt ihre besondere Perfidie. Die Grenze zwischen authentischem Material und Simulation wird systematisch porös.

Für Redaktionen bedeutet das: Verifikation darf nicht länger als nachgelagerter Prüfschritt verstanden werden, der unter Zeitdruck möglichst schnell ‚mitläuft‘. Sie muss integraler Bestandteil redaktioneller Prozesse werden. Die zentrale Frage lautet nicht mehr nur: Stimmt diese Information? Sondern: *Unter welchen technischen, visuellen und datenbezogenen Bedingungen ist sie entstanden?* KI verschiebt diese Beweislast: Nicht mehr die Fälschung muss außergewöhnlich gut sein – die Prüfung muss außergewöhnlich präzise werden.

Damit wächst der Bedarf an forensischer Kompetenz auf mehreren Ebenen zugleich: technisch, organisatorisch und edukativ. Es reicht nicht aus, einzelne Tools oder Verfahren zu beherrschen. Redaktionen müssen *verstehen*, wie KI-generierte Inhalte gemacht und in Umlauf gebracht werden, wo ihre Schwächen liegen und welche neuen Manipulationslogiken sich daraus ergeben. Forensik wird damit zur Voraussetzung journalistischer Arbeit im KI-Zeitalter.

Ein zentraler Baustein ist der systematische Einsatz von OSINT-Methoden (Open Source Intelligence) und Verifikations-Toolkits. Dazu gehören klassische Verfahren wie *Reverse Image Search*, Metadatenanalyse oder Quellenhistorien ebenso wie neuere Instrumente zur Erkennung synthetischer Inhalte. *Deepfake Detection*, Audio-Forensik oder Bildartefakt-Analysen sind keine Spezialthemen mehr für Ausnahmefälle, sondern zunehmend Teil des Alltagsgeschäfts.

Entscheidend ist jedoch nicht das einzelne Werkzeug, sondern seine institutionelle Verankerung. Verifikation darf nicht an einzelne Spezialist:innen delegiert werden, während der Rest der Redaktion unter Produktionsdruck darauf verzichtet. Sie muss als *kollektive Kompetenz* verstanden werden. Nur so lässt sich verhindern, dass forensisches Wissen lediglich punktuell vorhanden, nicht aber alltagstauglich ist.

Hinzu kommt eine zweite Ebene, die in vielen Redaktionen bislang unterbelichtet ist: die proaktive Auseinandersetzung mit Manipulationsszenarien. Einige Medienhäuser experimentieren bereits mit so genannten *Red Teams*, die gezielt Desinformationskampagnen, Deepfake-Szenarien oder KI-Missbrauch simulieren.¹²⁸ Solche Übungen sind mehr als technische Tests oder Sicherheitsexperimente: Sie schärfen das organisationale Problembewusstsein und machen sichtbar, wo Prozesse anfällig sind, wo Zuständigkeiten unklar bleiben und wo redaktionelle Routinen ausgenutzt werden könnten.

Gleichzeitig berührt Forensik eine *kommunikative* Dimension: In einer Öffentlichkeit, die zunehmend von Unsicherheit geprägt ist, reicht es nicht aus, Fälschungen intern zu erkennen. Redaktionen müssen ihre Prüfprozesse auch nach außen hin erklärbar machen. Transparente Verifikation, nachvollziehbare Korrekturen und offene Kommunikation über Unsicherheiten werden selbst zu Vertrauenssignalen. Forensik ist damit nicht nur ein Schutzschild gegen Desinformation, sondern Bestandteil journalistischer Beziehungsarbeit.

Der Resilienz-Effekt dieses Handlungsfelds liegt daher nicht in der Illusion vollständiger Sicherheit als in der Fähigkeit, mit Unsicherheit professionell und agil umzugehen. Redaktionen, die forensische Kompetenz systematisch aufbauen, stärken ihre Abwehrfähigkeit gegen Manipulation – und zugleich ihre Glaubwürdigkeit als Institutionen, die

Wahrheit nicht behaupten, sondern begründen. In einer KI-geprägten Öffentlichkeit wird Verifikation damit zur Kernkompetenz demokratischer Öffentlichkeitspflege.

Internationale Best Practices – etwa spezialisierte Verifikationseinheiten wie *AFP Fact Check* oder *BBC Verify* – zeigen, dass einzelne Werkzeuge allein nicht ausreichen. Der Stand von Forschung und Praxis verweist vielmehr auf die Kombination aus OSINT-Methoden, technischer Deepfake-Erkennung und simulationsbasierten *Red-Team*-Ansätzen. Genau diese Kombination entspricht auch den Empfehlungen aktueller Desinformations- und Sicherheitsforschung.

Hier greift das Framework systematisch ein. Forensik wird nicht als Spezialdisziplin für Ausnahmefälle behandelt, sondern als *organisatorische Infrastruktur*. Sie reagiert direkt auf jene strukturellen Brüche, die in den vorherigen Kapiteln beschrieben wurden: den Übergang von offensichtlicher Fälschung zu plausibler Simulation, von klaren Lügen zu probabilistisch erzeugten Wirklichkeiten.

Im Zusammenspiel mit Transparenz, *Human-in-the-Loop* und Ethik wird Forensik zu einem stabilisierenden Element redaktioneller KI-Resilienz. Sie sichert nicht nur Inhalte ab, sondern schützt die Fähigkeit des Journalismus, Öffentlichkeit als Raum recherchierter und überprüfter Wahrheit aufrechtzuerhalten – auch unter Bedingungen technologischer Unsicherheit.

Handlungsfeld 5: Ethik & Accountability:

Warum Verantwortung im KI-Zeitalter neu verankert werden muss

Ethik im Journalismus war nie verhandelbar. Sie war stets gebunden an konkrete Personen, Rollen und Entscheidungen: *Wer recherchiert? Wer entscheidet über Veröffentlichung? Wer korrigiert Fehler – und wer steht öffentlich dafür ein?* Diese klare Zuordenbarkeit von Verantwortung gehörte lange zu den stillen Voraussetzungen journalistischer Glaubwürdigkeit. Mit dem Einsatz von KI gerät genau dieses Gefüge unter Druck.

Wenn KI-Systeme Themen vorschlagen, Informationen gewichten, Texte formulieren, Sprache glätten, entstehen Entscheidungen nicht mehr ausschließlich im bewussten Urteil einzelner Journalist:innen, sondern werden technisch vorbereitet. Verantwortung wird dadurch tendenziell *diffus*. Und in dieser Diffusion steckt eines der größten Risiken für die *journalistische Integrität unter KI-Vorzeichen*.

Viele aktuelle KI-Leitlinien reagieren darauf mit ethischen Grundsätzen: Fairness, Transparenz, Nichtdiskriminierung, menschliche Aufsicht. Diese Prinzipien sind wichtig, aber sie bleiben oft folgenlos, solange sie nicht institutionell verankert werden. Die zentrale

Frage lautet daher nicht, welche Werte vertreten werden, sondern: *Wer trägt wann wofür Verantwortung – auch dann, wenn die KI beteiligt ist?* Genau an diesem Punkt setzt das Resilienz-Framework an: Es verschiebt den Fokus von abstrakter Ethik hin zu *Accountability* – also zu überprüfbarer, sichtbarer Verantwortung innerhalb der Organisation.

Ein zentraler Hebel dafür ist die explizite Benennung von Zuständigkeiten. Einige Redaktionen experimentieren bereits mit neuen Rollenmodellen – etwa KI-Editor:innen, Review-Verantwortlichen oder internen Governance-Gremien. Entscheidend ist dabei nicht die Jobbezeichnung, sondern die dahinterliegende Logik: Verantwortung darf nicht implizit bleiben. Sie muss *benannt, dokumentiert und überprüfbar* sein. Nur so lässt sich verhindern, dass sich Entscheidungen hinter Technik, Prozessen oder kollektiver Unschärfe verstecken.

Eng damit verbunden ist die Frage der Fehlerkultur. KI-Systeme produzieren keine Fehler im klassischen Sinn. Sie liefern Ergebnisse, die statistisch plausibel, aber faktisch falsch sein können. Wenn solche Fehler auftreten, entsteht schnell eine neue *Das-hat-die-KI-gemacht*-Ausrede. Diese schleichende Verschiebung ist strukturell gefährlich. Sie entlastet kurzfristig, untergräbt aber langfristig die Grundlage journalistischer Glaubwürdigkeit: die Bereitschaft, für eigene Entscheidungen einzustehen.

Resiliente Redaktionen brauchen deshalb explizite Regeln für den Umgang mit KI-Fehlern. *Wie werden sie erkannt? Wie werden sie korrigiert? Und wie werden sie kommuniziert?* Transparente Korrekturmechanismen, klare Eskalationspfade und offene Fehlerkommunikation sind in diesem Kontext kein Makel, sondern ein Ausdruck institutioneller Stärke. Sie zeigen dem Publikum: Verantwortung wird nicht delegiert – weder an Maschinen noch an anonymisierte Prozesse.

Ethik im KI-Zeitalter bedeutet darüber hinaus, Entscheidungen nicht nur retrospektiv zu rechtfertigen, sondern prospektiv zu reflektieren. Interne Ethik-Boards, Redaktionskonferenzen zu Grenzfällen oder regelmäßige Reflexionsformate schaffen Räume, in denen Unsicherheit zirkulieren darf. Sie ermöglichen es, Spannungen frühzeitig zu erkennen: *Wo verzerrt KI journalistische Perspektiven? Wo verstärkt sie Machtasymmetrien? Wo kollidiert Effizienz mit öffentlicher Verantwortung?*

Der Resilienz-Effekt dieses Handlungsfelds liegt in der Rückbindung von Technik an Verantwortung. Redaktionen, die Ethik und Accountability systematisch verankern, stärken nicht nur ihre interne Klarheit, sondern auch ihre öffentliche Legitimität. Sie machen sichtbar, dass KI-Einsatz kein Selbstzweck ist, sondern Teil einer professionellen, rechen-schaftspflichtigen Praxis. In einer Zeit, in der immer mehr Entscheidungen technisch vorbereitet werden, wird diese Sichtbarkeit von Verantwortung zur zentralen Währung des Vertrauens.

Handlungsfeld 6: Kompetenzaufbau, Weiterbildung und Organisationslernen: Warum KI-Resilienz ohne Lernen nicht zu haben ist

KI-Resilienz ist kein Zustand, den Redaktionen einmal erreichen und anschließend abhaken können. Sie ist ein fortlaufender Prozess. Und dieser Prozess ist ohne systematischen Kompetenzaufbau nicht denkbar. Gerade hier zeigt sich eine der größten Schwächen vieler gegenwärtiger KI-Strategien im Medienbereich: Sie setzen auf Tools, Pilotprojekte und punktuelle Trainings, ohne Lernen als dauerhafte redaktionelle Kernaufgabe zu projektieren.

Die Ausgangslage ist dabei paradox. Einerseits wächst der Erwartungsdruck, KI produktiv einzusetzen, rapide – aus ökonomischen Gründen, aus Effizienzerwägungen, aus Konkurrenzangst. Andererseits fühlen sich viele Journalist:innen unzureichend vorbereitet. Die Folge ist ein gefährliches Kompetenzgefälle: Einige wenige werden zu informellen KI-Expert:innen, während andere KI entweder meiden oder unkritisch nutzen. Beides untergräbt die organisationale Resilienz. Denn resiliente Organisationen zeichnen sich durch *verteilte, anschlussfähige Kompetenz* aus.

Kompetenzaufbau im KI-Zeitalter darf deshalb nicht auf Tool-Einweisungen reduziert werden. Es reicht nicht, zu erklären, wie man Prompts formuliert oder welche Software wofür geeignet ist. Entscheidend ist ein grundlegendes Verständnis davon, *wie KI-Systeme funktionieren*, wo ihre Grenzen liegen und welche strukturellen Risiken mit ihrem Einsatz verbunden sind. Wer nicht versteht, dass KI probabilistisch arbeitet, dass sie keine Wahrheit kennt und dass sie systematisch verzerren kann, wird ihre Ergebnisse entweder überschätzen – oder reflexhaft ablehnen.

Für Redaktionen bedeutet das: KI-Kompetenz ist heute Teil journalistischer Grundbildung. Ähnlich wie Recherchemethoden, Quellenkritik oder Presserecht muss auch KI-Resilienz strukturell verankert werden. Dazu gehören Basismodule, die technisches Grundverständnis, rechtliche Rahmenbedingungen – etwa den EU AI Act – sowie ethische und organisationale Fragen miteinander verbinden. Ziel ist ausdrücklich nicht, Journalist:innen zu Informatiker:innen zu machen (obwohl es diesen Ansatz in einigen journalistischen Ausbildungen schon gibt oder gegeben hat), sondern zu *mündigen Anwender:innen mit Urteilskraft*.

Ein zweiter zentraler Baustein ist die Organisation von Lernräumen. KI verändert sich schnell, feste Curricula veralten rasch. Deshalb braucht es Formate, die experimentelles, iteratives Lernen ermöglichen: interne KI-Wikis, *Learning Sprints*, Redaktionslabs oder temporäre Pilotprojekte mit klarer Auswertungslogik. Entscheidend ist dabei weniger das einzelne Format als die institutionelle Absicherung: Lernen darf nicht als Zusatzbelastung ‚on top‘ organisiert werden, sondern muss Teil der regulären Arbeitszeit sein.

Ein weiterer kritischer Punkt ist die Übersetzung individuellen Lernens in organisationale Kompetenz. Schulungen entfalten nur dann nachhaltige Wirkung, wenn Wissen dokumentiert, geteilt und aktualisiert wird. Gute Praxis entsteht dort, wo Lernerfahrungen nicht an einzelne Personen gebunden bleiben, sondern in Prozesse, Leitlinien und Routinen überführt werden. Andernfalls entsteht Abhängigkeit von Einzelpersonen – ein klassisches Risiko organisationaler Fragilität.

Organisationslernen bedeutet in diesem Zusammenhang auch, Fehler als Lernanlässe zu begreifen. KI-Experimente werden scheitern. Tools werden falsche Ergebnisse liefern. Prozesse werden angepasst werden müssen. Entscheidend ist nicht, ob Fehler passieren, sondern *wie Redaktionen mit ihnen umgehen*. Lernende Organisationen schaffen Feedbackschleifen, Reflexionsräume und Möglichkeiten zur Korrektur – ohne Schuldzuweisungen, aber mit klarer Verantwortlichkeit.

Der Resilienz-Effekt dieses Handlungsfelds liegt auf mehreren Ebenen. Kontinuierlicher Kompetenzaufbau reduziert Abhängigkeiten von externen Anbietern, verhindert interne Spaltungen und stärkt die Fähigkeit von Redaktionen, KI souverän einzusetzen – auch unter Zeitdruck. Vor allem wirkt er präventiv gegen Überforderung, Unsicherheit, Ängste und Abwehrreaktionen. KI-Resilienz bedeutet in diesem Sinne also: lernfähig bleiben. Nicht nur individuell, sondern als Organisation. Redaktionen, die Lernen als strategische Ressource begreifen, erhöhen ihre Anpassungsfähigkeit – und sichern damit ihre Handlungsfähigkeit in einem Umfeld, das sich schneller verändert als jede einzelne Technologie.

Handlungsfeld 7: Publikumsdialog & Beziehungsarbeit:

Warum KI-Resilienz ohne Beziehung nicht funktioniert

Während die vorherigen Handlungsfelder vor allem auf interne Strukturen zielen, öffnet sich hier der Blick nach außen. KI-Resilienz entscheidet sich nicht allein im Maschinenraum der Redaktion, sondern im Verhältnis zum Publikum. Denn Journalismus existiert nicht für sich selbst. Seine Legitimität entsteht dort, wo Menschen ihm Vertrauen schenken, ihn verstehen – oder ihm widersprechen können. Genau dieses Verhältnis wird durch KI tiefgreifend verändert.

Für das Publikum ist KI meist unsichtbar präsent. Texte erscheinen schneller, Antworten wirken glatter, Zusammenfassungen präziser, manchmal auch austauschbarer. Gleichzeitig wächst die Unsicherheit: Wer spricht hier eigentlich? Ein Mensch, ein System oder eine hybride Konstellation? Wird KI eingesetzt, um zu helfen – oder um Arbeit zu verstecken? Schweigen Redaktionen zu diesen Fragen, entsteht Misstrauen. Erklären sie ihren

Umgang mit KI, entsteht Anschlussfähigkeit. Publikumsdialog wird damit nicht zur Kür, sondern zur Voraussetzung journalistischer Glaubwürdigkeit.

Entscheidend ist, Dialog nicht mit Transparenz zu verwechseln. Transparenz informiert, Dialog verbindet. Beziehung entsteht dort, wo Redaktionen bereit sind, ihren Einsatz von KI zu erklären, Kritik aufzunehmen und Unsicherheiten auszuhalten. Gerade im KI-Kontext ist das entscheidend: Viele Fragen lassen sich nicht abschließend beantworten. Resilienz entsteht deshalb auch nicht durch perfekte Antworten, sondern durch *Gesprächsbereitschaft*.

Konkret heißt das: KI-Einsatz muss explizit Teil redaktioneller Dialogformate werden. *Ask-Me-Anything*-Sessions, Redaktionsgespräche, Publikumskonferenzen oder Newsletter können genutzt werden, um offenzulegen, wo KI hilft, wo sie begrenzt ist und wo Menschen entscheiden. Nicht jede Redaktion braucht neue Formate; oft reicht es, bestehende redaktionelle Beteiligungsformate um die KI-Frage zu erweitern. Entscheidend ist, dass KI nicht als technisches Detail behandelt wird, sondern als Thema bewusster öffentlicher Aushandlung.

Gleichzeitig eröffnet KI neue Möglichkeiten partizipativer Formate. Community-orientierte KI-Projekte – etwa bei Datenrecherchen, Themenclustern oder der Auswertung großer Dokumentensätze – können genutzt werden, um Publikum aktiv einzubeziehen. Wichtig ist dabei die Gestaltung: KI darf nicht zum Ersatz für Beteiligung werden, sondern zum Anlass für sie. Resiliente Redaktionen nutzen Technologie, um Dialog zu vertiefen – nicht, um Distinktion oder Distanz zu begründen.

Ein besonders sensibler Punkt ist der Umgang mit Publikumsfeedback. KI kann helfen, Rückmeldungen zu sortieren, Muster zu erkennen oder Themen zu priorisieren. Doch Automatisierung darf hier nicht das Zuhören ersetzen. Beziehung ist kein Datensatz, sondern ein Prozess. Wer Feedback nur noch algorithmisch verarbeitet, riskiert, die emotionale Dimension von Öffentlichkeit zu verlieren – und damit genau jene Vertrauensbasis, auf die Journalismus angewiesen ist.

Der Resilienz-Effekt dieses Handlungsfelds liegt in der Rückbindung journalistischer Arbeit an ihre soziale Funktion. Redaktionen, die ihren KI-Einsatz erklärbar machen und Dialog aktiv gestalten, schaffen eine Pufferzone gegen Vertrauensverluste und stärken im gleichen Schritt die Medienkompetenz des Publikums.

Hier zeigt sich die größte Leerstelle bestehender KI-Leitlinien. Während Transparenz vielfach adressiert wird, bleibt der systematische Dialog mit dem Publikum meist außen vor – oder wird nicht als Gegengewicht zur technischen Hypertrophie verstanden. Nur wenige Ausnahmen – etwa erklärende KI-Formate im KI-Kontext von Q&A-Formaten mit

Leserfragen wie bei „LeMonde“, einem KI-gestützten Publikums-Interface wie bei der *Washington Post* oder bei eher traditionellen dialogischen Redaktionsansätzen wie bei „ProPublica“ oder beim „Guardian“¹²⁹ – deuten an, welches Potenzial hier gehoben werden kann. Indem dieses Framework jedoch den Publikumsdialog als eigenständiges Handlungsfeld definiert, verlängert es KI-Resilienz konsequent über die Redaktion hinaus in die Öffentlichkeit. KI wird nicht nur erklärt, sondern gemeinsam mit den Mediennutzen verhandelt. Genau darin liegt seine demokratische Qualität.

Die Stärke dieses Frameworks zeigt sich gerade hier besonders deutlich: KI-Resilienz bedeutet dann nicht nur, Technik zu kontrollieren, sondern Vertrauen aktiv zu pflegen. Nicht durch Inszenierung oder Perfektion, sondern durch Gesprächsbereitschaft, Erklärbarkeit und Präsenz. In einer zunehmend synthetischen Öffentlichkeit wird genau diese Fähigkeit der Beziehungspflege zur knappen Ressource – und in der Vertrauensökonomie damit zur eigentlichen Zukunftskompetenz des Journalismus.

In der Zusammenschau wird deutlich: Die hier entwickelten Handlungsfelder bilden kein loses Nebeneinander einzelner Maßnahmen, sondern ein integriertes Modell. Sie ordnen fragmentierte Leitlinien neu, schließen institutionelle Lücken und machen KI-Resilienz erstmals als organisationsweite Kernkompetenz sichtbar.

Die folgenden Praxisbeispiele aus internationalen Redaktionen illustrieren daher nicht nur gelungene Ansätze, sondern auch typische Bruchstellen und Warnsignale bei der KI-Integration. Sie machen sichtbar, wo einzelne Handlungsfelder bereits greifen – und wo fehlende Verzahnungen zu Vertrauensverlust, Verantwortungsdiffusion oder institutionellen Krisen führt. Gerade im Vergleich von Best Practices und Fehlentwicklungen zeigt sich, warum KI-Resilienz keine Frage des guten Willens ist, sondern eine Frage organisatorischer Gestaltung.

6.3 Das Framework im Praxistest: Best Practices und Warnzeichen

Die zuvor beschriebenen sieben Handlungsfelder bilden den Kern des KI-Resilienz-Frameworks. In Kapitel 3 wurde gezeigt, wie sich entlang von zwei Achsen – aktiver KI-Gestaltung und redaktioneller Souveränität – idealtypische Muster journalistischer KI-Resilienz ausbilden. Dort stand die *Organisation als Ganze* im Mittelpunkt:

governance-orientiert, kompetenzgetrieben, ökonomisch motiviert, vorsorgeorientiert oder experimentell.

Dieses Kapitel verfolgt eine andere Perspektive. Es fragt nicht, welches Muster ein Medienhaus insgesamt verkörpert, sondern wie gut einzelne Dimensionen des Frameworks in der Praxis ausgeprägt sind – und was passiert, wenn zentrale Handlungsfelder fehlen oder schlecht verzahnt sind. Das Ziel ist ein *Praxistest des Frameworks*: Jeder der sieben Bausteine wird mit konkreten Beispielen unterlegt – dort, wo er getragen wird, und dort, wo Leerräume sichtbar werden.

Die Beispiele sind dabei bewusst *selektiv und kontrastierend* gewählt. Sie bewerten keine Häuser ‚insgesamt‘, sondern fokussieren auf einzelne Handlungsfelder. Das bedeutet: Ein und dieselbe Redaktion kann in einem Feld als Vorbild gelten und in einem anderen deutliche Defizite zeigen. Gerade diese Fragmentierung ist Ausdruck der aktuellen Übergangsphase: Viele Häuser haben bereits KI-Policies, einzelne Tools oder Teams etabliert, aber selten ein konsistent verzahntes System. Das KI-Resilienz-Framework dient in diesem Kapitel als Vergrößerungsglas, um *Best Practices* und *Warnzeichen* zu identifizieren – und zu zeigen, wo aus Insellösungen organisationale Resilienz wird und wo nicht.

Handlungsfeld 1: Transparenz & Kennzeichnung – Sichtbarkeit von KI als Vertrauensfrage

Fokus des Frameworks: Dieses Feld adressiert die Frage, wie sichtbar der Einsatz von KI für das Publikum ist – und wie zuverlässig sich Nutzer:innen darauf verlassen können, dass sie über Rolle und Grenzen von KI informiert werden. Transparenz meint hier nicht technische Detailoffenlegung, sondern *Erwartungssicherheit*: Wann war KI beteiligt, wofür diente sie, und wer trägt Verantwortung?

Best Practices: Sichtbare Selbstverpflichtung: Die BBC hat mit „BBC Verify“ nicht nur eine Verifikationseinheit aufgebaut, sondern macht deren Arbeit explizit sichtbar. In der offiziellen Ankündigung wird „Verify“ als Markenzeichen für Transparenz beschrieben: Das Team aus forensisch arbeitenden Journalist:innen soll offenlegen, wie Fakten geprüft, Bilder verifiziert und Desinformation zurückgewiesen werden. Transparenz richtet sich hier nicht nur auf Inhalte, sondern auf Prüfprozesse – ein zentrales Element des Frameworks.¹³⁰

Der „Guardian“ hat 2023 seinen Editorial Code aktualisiert und darin explizite Leitlinien zur Nutzung generativer KI verankert. Der Code hält fest, dass KI-Systeme grundsätzlich nicht dazu genutzt werden sollen, Texte oder Bilder zu erzeugen, die direkt in die Berichterstattung eingehen, außer in eng definierten Ausnahmefällen. Solche „exceptional

use“-Fälle müssen von einer verantwortlichen Leitungsperson genehmigt und gegenüber Leser:innen klar gekennzeichnet werden. Damit wird Transparenz zur *verbindlichen Pflicht*, nicht zur nachträglichen Option.¹³¹

Auch das „Schweizer Radio und Fernsehen“ (SRF) hat Transparenz zum Prinzip gemacht, indem es ein eigenes KI-Kapitel in seine publizistischen Leitlinien aufgenommen hat. Die Leitlinien orientieren sich an drei Grundprinzipien der SRG – Verantwortung, Transparenz, Vertraulichkeit – und formulieren spezifische Handlungsanweisungen zur Nutzung von KI-Anwendungen sowie KI-generierter oder -bearbeiteter Texte, Bilder und Audios. KI wird damit nicht „informell“ geregelt, sondern in den normativen Kern des Hauses integriert.¹³²

Warnzeichen: *Versteckter KI-Einsatz und reaktive Offenlegung:* Die Auseinandersetzung um „Sports Illustrated“ zeigt, welche Folgen intransparenter KI-Einsatz haben kann: Recherchen des Magazins „Futurism“ legten 2023 nahe, dass auf der Website Produkttexte unter erfundenen Autor:innen publiziert wurden, inklusive offenbar KI-generierter Porträtfotos; die Inhalte stammten von einem externen Dienstleister. Nach öffentlicher Kritik wurden entsprechende Beiträge entfernt, der Dienstleister gewechselt und später der CEO des Mutterunternehmens The Arena Group entlassen. Entscheidend für das Framework: Das Problem war nicht nur der mögliche Einsatz von KI, sondern die *bewusste Intransparenz* gegenüber dem Publikum.¹³³

Bei CNET wurde zwischen 2022 und 2023 ein internes KI-System genutzt, um Finanzratgeber-Artikel zu verfassen. Erst nachdem externe Beobachter:innen Fehler entdeckten, wurde offen kommuniziert, dass hier eine „internally designed AI engine“¹³⁴ im Einsatz war. Eine nachträgliche Auswertung ergab, dass 41 von 77 entsprechenden Artikeln Korrekturen benötigten.¹³⁵ Die Korrekturvermerke erschienen erst, nachdem die Praxis öffentlich problematisiert worden war. Transparenz fungierte damit als *Schadensbegrenzung*, nicht als präventive Praxis.

Fazit: In beiden Positivbeispielen wird Transparenz institutionell verankert: durch Codes, Marken („BBC Verify“) und Leitlinien. Sie schafft einen *stabilen Erwartungshorizont* für das Publikum und für Mitarbeitende. In den Negativfällen wird KI-Einsatz erst dann sichtbar, wenn externe Kritik dies erzwingt. Das Framework macht hier deutlich: Fehlende Transparenz ist nicht bloß ein Kommunikationsdefizit, sondern unterminiert die Grundlage aller weiteren Handlungsfelder – insbesondere Ethik, Human-in-the-Loop und Publikumsbeziehung.

Handlungsfeld 2: Human-in-the-Loop & Verantwortungszuordnung – Wer entscheidet im Zweifel?

Fokus des Frameworks: Human-in-the-Loop meint im Framework nicht nur einen letzten Kontrollblick, sondern die *klare Zuordnung von Verantwortung* entlang des Produktionsprozesses: Wer prüft KI-Outputs wann, mit welcher Autorität und haftet für Fehler?

Best Practices: *Klare Rollen und Grenzen:* Die „New York Times“ kommuniziert – in öffentlichen Statements und internen Standards – konsequent, dass KI als Werkzeug eingesetzt werden kann, die letztendliche Verantwortung aber bei Journalist:innen bleibt. Experimentelle Projekte (etwa bei Datenanalyse oder Formatentwicklung) werden in dedizierten Teams angesiedelt; KI-Outputs gelten als Ausgangsmaterial, nicht als fertige Produkte. Human-in-the-Loop ist hier organisatorisch eingebaut, nicht nur postuliert.¹³⁶

Die BBC koppelt KI-Einsatz systematisch an redaktionelle Verantwortung. In Darstellungen ihrer AI-Richtlinien wird betont, dass jede Nutzung an den publizistischen Leitlinien und Kernwerten – insbesondere Genauigkeit, Fairness und Unabhängigkeit – gespiegelt werden muss und stets eine verantwortliche redaktionelle Instanz benannt ist.¹³⁷ Auch „BBC Verify“ ist als *Service in der Verantwortung der Redaktion* angelegt, nicht als ‚automatisierte Wahrheitsmaschine‘.

Die „Associated Press“ hat im August 2023 Leitlinien veröffentlicht, die klar festhalten, dass generative KI nicht zur Erstellung veröffentlichter Texte, Fotos, Videos oder Audios verwendet werden darf. Outputs aus Tools wie ChatGPT werden als „unvetted source material“ behandelt,¹³⁸ das redaktionell zu prüfen ist; zugleich wird Mitarbeitenden explizit abgeraten, vertrauliche oder sensible Informationen in solche Systeme einzugeben.¹³⁹ Damit wird Human-in-the-Loop zur *Nicht-Delegation von Verantwortung*: KI darf unterstützen, aber nicht publizistisch entscheiden.

Warnzeichen: *De-facto-Automatisierung ohne tragfähige Rollen:* Beim Technikportal CNET zeigte sich, wie Human-in-the-Loop formal behauptet, aber praktisch ausgehöhlt werden kann. Offiziell sollten alle KI-Texte von Redakteur:innen gegengelesen werden; faktisch blieben schwerwiegende Fehler – etwa falsche Zinsberechnungen und verwechselte Finanzbegriffe – über Wochen online. Erst nachdem externe Medien die Probleme aufgegriffen hatten, wurden systematische Korrekturen eingeleitet.¹⁴⁰ Die Kontrollinstanz existierte auf dem Papier, hatte aber im Produktionsalltag offenbar nicht genügend Gewicht.

Bei „Gizmodo“ (Rubrik io9) wurde ein KI-generierter Artikel über die Chronologie von Star-Wars-Filmen und -Serien mit zahlreichen Fehlern veröffentlicht. Das Stück erschien unter dem Label „Gizmodo Bot“, ohne dass die zuständige Redaktion eigener Aussage

zufolge in Erstellung oder Freigabe eingebunden war.¹⁴¹ Hier fungierte KI zeitweise als *Parallelredaktion ohne echte Human-in-the-Loop-Struktur*: Die Verantwortung war organisatorisch entkoppelt.

Fazit: Die Beispiele zeigen: Human-in-the-Loop ist nur dann ein resilientes Schutzinstrument, wenn Rollen und Befugnisse *klar definiert und im Alltag wirksam* sind. Wo Verantwortung aus Bequemlichkeit, Kostendruck oder Tool-Euphorie an Systeme ‚mitwandert‘, entstehen neue Fehlerklassen – und eine gefährliche Grauzone der Verantwortungsdiffusion. Das Framework macht hier deutlich, dass Human-in-the-Loop strukturell – über Rollen, Freigaberoutinen, Eskalationspfade – zu verankern ist, nicht nur rhetorisch.

Handlungsfeld 3: Quellen- & Datenvertrauen – Digitale Souveränität bei sensiblen Informationen

Fokus des Frameworks: Dieses Feld übersetzt klassischen Quellenschutz in KI-Infrastrukturen. Zentral ist die Frage, *welche Daten in welche Systeme gelangen* und wie sich Redaktionen vor unkontrollierbaren Abhängigkeiten schützen.

Best Practices: *Bewusste Begrenzung und geschützte Räume:* Die Associated Press koppelt ihre generellen KI-Leitlinien explizit an Datenschutz und Vertraulichkeit. Mitarbeitende werden angehalten, keine KI-generierten Inhalte zu veröffentlichen, Outputs nur als ungeprüftes Rohmaterial zu nutzen und vertrauliche Daten zu schützen.¹⁴² Praktisch heißt das: Für investigatives Material, vertrauliche Dokumente oder Identitäten von Hinweisgeber:innen werden cloudbasierte Gen-KI-Dienste nicht genutzt. Digitale Souveränität wird durch *Selbstbegrenzung gesichert*.

Das SRF nutzt KI nach eigenen Angaben unter anderem für Transkription und technische Unterstützung, verankert die Grundprinzipien Verantwortung, Transparenz und Vertraulichkeit aber explizit in seinen Leitlinien und Handlungsanweisungen zur Nutzung von KI-Anwendungen sowie KI-generierten oder -bearbeiteten Inhalten.¹⁴³ Damit sind die sensibelsten Punkte – Umgang mit Material, das Rückschlüsse auf Quellen zulässt – normativ adressiert, selbst wenn nicht jede Einzelsituation detailliert geregelt ist.

Das Transkriptions-Tool „Good Tape“, das vom dänischen Medium „Zetland“ hervorgegangen ist, illustriert eine weitere Form von Quellen- und Datenvertrauen. Der Dienst verspricht explizit, Audio-Dateien verschlüsselt auf europäischen Servern zu speichern, keine Trainingsdaten daraus abzuleiten und keine Daten mit Dritten zu teilen.¹⁴⁴ Für Redaktionen, die auf automatisierte Transkription angewiesen sind, schafft dies einen

abgeschotteten Datenraum, der deutlich näher an klassischen Quellenschutz-Standards liegt als generische Cloud-Services.

Warnzeichen: *Strukturelle Abhängigkeiten und agentische KI:* Das Netzwerk „Ippen Digital“ experimentiert offensiv mit KI-gestützter Skalierung von Inhalten. Berichte vom European Publishing Congress und Fachnewsletter zeigen, wie tausende lokal zugeschnittene Wahlergebnis-Artikel mit Hilfe großer Sprachmodelle generiert wurden – auf Basis eines sehr umfangreichen Prompts und unter starker Automatisierung. Parallel dazu kommuniziert Ippen Digital Pläne, „agentische KI“ einzusetzen,¹⁴⁵ die komplexe Arbeitsabläufe weitgehend selbständig ausführt. Öffentlich einsehbare Materialien konzentrieren sich auf Effizienz und Reichweite; Fragen nach *Trennung sensibler von nicht sensiblen Daten*, On-Prem-Lösungen oder klaren Sperrzonen für vertrauliche Inhalte werden bislang kaum adressiert.¹⁴⁶

Auch beim Einsatz externer Dienste – etwa cloudbasierter Transkriptionstools, automatischer Übersetzung oder KI-gestützter Textanalyse – ist aus öffentlich zugänglichen Informationen oft nicht klar, welche Daten wohin fließen und unter welchen Bedingungen sie weiterverarbeitet werden. Das ist weniger ein spezifischer Vorwurf an ein Haus als ein strukturelles Muster: *Cloud-Abhängigkeit ohne ausreichende Governance*.

Fazit: Quellen- & Datenvertrauen ist im Framework das Handlungsfeld, das besonders stark von externen Infrastrukturen geprägt ist. Best-Practice-Beispiele zeigen, dass sich Risiken durch klare Verbote (AP) und prinzipienbasierte Leitlinien (SRF) zumindest begrenzen lassen. Tools wie „Good Tape“ demonstrieren, dass es alternative, quellenverträglichere Lösungen geben kann. Wo hingegen agentische KI und großskalige Automatisierung vorangetrieben werden, ohne Datenflüsse transparent und differenziert zu regeln, entstehen *strukturelle Verwundbarkeiten*, die sich durch andere Handlungsfelder (z.B. Ethik oder Transparenz) nur begrenzt kompensieren lassen.

Handlungsfeld 4: Forensik & Verifikation – Wahrheitsprüfung als Infrastruktur

Fokus des Frameworks: Dieses Feld verankert Verifikation nicht als individuellen Reflex, sondern als *institutionelle Infrastruktur*: spezialisierte Teams, Tools und Routinen zur Erkennung synthetischer Inhalte, Deepfakes und manipulierten Materials.

Best Practices: *Institutionalisierte Prüfkompetenz:* Dadurch, dass die BBC 2023 mit „BBC Verify“ eine eigenständige Marke geschaffen hat, die forensische Recherche, O-SINT-Methoden, Datenanalyse und Deepfake-Prüfung bündelt, arbeitet das Team

sichtbar über Formate und Plattformen hinweg; seine Existenz signalisiert: Verifikation ist ein eigener, ressourcenstarker Arbeitsbereich, kein beiläufiger Prüfschritt.

Internationale Agenturen wie „Reuters“ und „Agence France-Presse“ (AFP) haben spezialisierte Fact-Checking-Teams aufgebaut, die systematisch mit OSINT-Toolkits, Verifikationssoftware und forensischer Bild- und Videoanalyse arbeiten. Berichte über den KI-Einsatz bei „Reuters“ beschreiben Pläne, automatisierte Unterstützung bei der Verifikation von Quellen und Informationen weiter auszubauen.¹⁴⁷ Damit wird forensische Arbeit schrittweise in die technische Infrastruktur eingebettet.

Die Kombination aus analogen Recherchetechniken, spezialisierten Teams und digitalen Tools findet sich auch bei Häusern wie der „New York Times“, deren Daten- und Visual-Investigations-Teams regelmäßig offenlegen, wie sie Bild-, Video- und Dokumentenmaterial prüfen. Solche Einheiten sind weniger an einzelne Technologien gebunden als an eine *Methodenkultur der Überprüfung*.

Warnzeichen: *Verifikation als nachgelagerte Formalie:* Der Fall CNET zeigt, dass fehlende systematische Verifikation unmittelbar zu Qualitätsproblemen führen kann. Die KI-generierten Finanzartikel wiesen nicht nur vereinzelt Tippfehler, sondern strukturelle Rechen- und Begriffsfehler auf, etwa zur Verzinsung von Sparprodukten oder zur Unterscheidung von APR (Effektiver Jahreszins/Annual Percentage Rate) und APY (Jährliche prozentuale Rendite/Annual Percentage Yield).¹⁴⁸ Erst eine nachträgliche Überprüfung ergab, dass mehr als die Hälfte der Artikel Korrekturen benötigte. Hier fehlte eine forensische Instanz, die KI-Outputs als *zu prüfendes Rohmaterial* behandelt.

Bei „Gizmodo“/„io9“ wurde der bereits erwähnte fehlerhafte Star-Wars-Artikel zunächst ohne Prüfung veröffentlicht und erst im Nachgang öffentlich als problematisch markiert. Weder existierte ein etabliertes Verifikationsteam für solche Inhalte, noch wurden Leser:innen proaktiv durch transparente Darstellung der Prüfschritte eingebunden.

Fazit: Forensik & Verifikation verschiebt im Framework die Perspektive von ‚Stimmt diese Information?‘ zu ‚Unter welchen Bedingungen ist sie entstanden und geprüft worden?‘. Institutionalisierte Einheiten wie „BBC Verify“ oder Fact-Checking-Teams schaffen eine *organisatorische Prüfkapazität*, die nicht allein von individueller Erfahrung abhängt. Fehlende Strukturen, wie in den Negativbeispielen, führen dazu, dass KI-Fehler erst durch externe Kritik sichtbar werden – ein klassisches Zeichen mangelnder Resilienz.

Handlungsfeld 5: Ethik & Accountability – Governance gegen Verantwortungsdiffusion

Fokus des Frameworks: Ethik & Accountability meint geregelte *Rechenschaftsmechanismen*: Wer entscheidet über KI-Einsatz, wie werden Konflikte bearbeitet, wie geht ein Haus mit Fehlern um?

Best Practices: *Ethik als sichtbare Selbstbindung*. Der im Jahr 2023 aktualisierte Editorial Code des „Guardian“ verknüpft KI-Einsatz mit klaren Verantwortlichkeiten. Generative KI soll nur in Ausnahmefällen Inhalte erzeugen, die direkt veröffentlicht werden. Solche Fälle müssen von einer verantwortlichen Führungsperson genehmigt und für Leser:innen gekennzeichnet werden. Ethik wird hier als *Governance-Instrument* operationalisiert, nicht als abstrakter Wert.

Die „New York Times“ arbeitet seit langem mit Standards- und Ethik-Teams, die redaktionelle Dilemmata begleiten und Korrekturprozesse verantworten. Öffentliche „Corrections“-Rubriken und interne Richtlinien zur Kennzeichnung und Korrektur von Fehlern machen Accountability *verfahrenstechnisch überprüfbar*. In Diskussionen um KI wird immer wieder betont, dass Verantwortung beim Menschen bleibt, auch wenn Tools Teile der Arbeit übernehmen.

Die jüngere Partnerschaft zwischen „Axios“ und OpenAI (Januar 2025) enthält eine klare Aufgabentrennung: OpenAI finanziert und unterstützt die Expansion lokaler Newsletter in vier Städten (Pittsburgh, Kansas City, Boulder, Huntsville), liefert technische Infrastruktur und ermöglicht die Nutzung von Axios-Inhalten in ChatGPT-Antworten. Zugleich betont „Axios“-CEO Jim VandeHei, dass KI nicht dazu genutzt werde, Nachrichtenartikel zu schreiben, sondern um eine Infrastruktur für die Erstellung, Distribution und Monetarisierung unseres Journalismus aufzubauen.¹⁴⁹ Die Vereinbarung zielt auf die Förderung von Qualitätsberichterstattung mit Fokus auf Lokaljournalismus, wie OpenAI proklamiert.¹⁵⁰ Die redaktionelle Verantwortung und die Erstellung der Inhalte soll gemäß der Vereinbarung explizit bei „Axios“-Journalist:innen verbleiben. Das ist kein umfassender Ethikkodex, zeigt aber: Externe Technologiepartnerschaften lassen sich so gestalten, dass Verantwortung explizit bei der Redaktion verortet bleibt. Das ist kein umfassender Ethikkodex, zeigt aber: Externe Technologiepartnerschaften lassen sich so gestalten, dass Verantwortung explizit bei der Redaktion verortet bleibt.

Warnzeichen: *Intransparente Entscheidungen und fehlende Aufarbeitung*. Im Fall „Sports Illustrated“ ging es nicht nur um KI-Einsatz an sich, sondern um Täuschung und fehlende Rechenschaft. Die Nutzung offenbar fiktiver Autor:innenprofile und KI-Porträts bei Produkttexten wurde erst nach externer Berichterstattung sichtbar; Reaktionen des

Mutterunternehmens blieben zunächst defensiv und verlagerten Verantwortung auf den Drittanbieter. Eine umfassende öffentliche Aufarbeitung, die Verantwortlichkeiten und Lernprozesse transparent gemacht hätte, blieb begrenzt.

Bei „Men’s Journal“, ebenfalls Teil der Arena Group, erschien 2023 ein KI-gestützter Gesundheitsartikel über Testosteron, der nach einer Überprüfung mindestens 18 gravierende fachliche Fehler enthielt. Die Korrekturen erfolgten leise; der ursprüngliche Umfang der Fehler wurde erst durch externe Analyse sichtbar.¹⁵¹ Die interne Disclosure-Formel, der Artikel sei von der Redaktion überprüft worden, stand in deutlichem Widerspruch zum tatsächlichen Qualitätsniveau – ein Beispiel für *unterlaufene Accountability-Versprechen*.

Fazit: Ethik & Accountability fungiert im Framework als *Scharnier* zwischen interner Governance und öffentlicher Legitimität. Die Positivbeispiele zeigen, dass klare Regeln, Zuständigkeiten und Korrekturmechanismen auch unter KI-Bedingungen Vertrauen stabilisieren können. Die Negativbeispiele machen deutlich, wie schnell Verantwortung diffundiert, wenn externe Partner, automatisierte Prozesse und wirtschaftlicher Druck zusammenkommen – und wie wichtig es ist, Fehler nicht nur zu korrigieren, sondern *institutionell zu bearbeiten*.

Handlungsfeld 6: Kompetenzaufbau & Training – Kollektive Urteilskraft statt Inselwissen

Fokus des Frameworks: Dieses Feld adressiert die *Lernfähigkeit der Organisation*. KI-Resilienz setzt voraus, dass nicht nur wenige Spezialist:innen, sondern ganze Redaktionen verstehen, was KI kann und wo ihre Grenzen liegen.

Best Practices: *Lernen als strukturierter Prozess:* Die „Financial Times“ hat mit dem internen „AI Playground“ einen geschützten Raum geschaffen, in dem Redakteur:innen mit großen Sprachmodellen auf Basis eigener Inhalte experimentieren können. Das Tool ist so gestaltet, dass Copy der FT nicht nach außen abfließt; zugleich können Journalist:innen Prompts testen, Ausgaben bewerten und Best-Practices teilen.¹⁵² Dies stärkt nicht nur Tool-Wissen, sondern verankert KI-Kompetenz in der Breite der Redaktion.

„Zetland“ aus Dänemark hat aus einem internen Problem – zeitraubende Transkription – sein eigenes KI-Tool entwickelt: „Good Tape“ wird inzwischen weit über die eigene Redaktion hinaus genutzt. Die Entstehungsgeschichte zeigt, wie praktische Experimente, Rückkopplungsschleifen mit Nutzer:innen und kontinuierliche Anpassung zu einem Lernprozess führen, der sowohl Produkt als auch Redaktion formt.¹⁵³ Entscheidend für das

Framework: KI wird hier als *Gegenstand gemeinsamen Lernens*, nicht nur als technisches Projekt verstanden.

Auch bei der „New York Times“ lässt sich beobachten, dass KI-Kompetenz über eigene Teams, Schulungsprogramme und experimentelle Projekte systematisch aufgebaut wird. Trainings, interne Wissensplattformen und Austauschformate tragen dazu bei, dass KI-Fragen nicht nur an eine Innovationsabteilung delegiert werden, sondern in der Fläche der Redaktion bearbeitet werden.

Warnzeichen: *Fragmentiertes Wissen und Abhängigkeit:* In vielen Häusern, insbesondere im Lokal- und Regionalbereich, berichten Journalist:innen von *individuellem Herumexperimentieren mit KI-Tools*, ohne dass klare Schulungsangebote oder gemeinsame Standards etabliert wären. Öffentliche Dokumente zu KI-Einsatz – etwa bei Unternehmen wie „Reach plc“, das mit dem hausinternen Tool „Guten“ journalistische Texte automatisiert auf verschiedene Markentile umschreibt –¹⁵⁴ konzentrieren sich stark auf Effizienzgewinne und Strukturveränderungen, während Fragen nach systematischem Kompetenzaufbau nur am Rand auftauchen. Gewerkschaftliche Stellungnahmen weisen darauf hin, dass Stellenabbau mit der Erwartung einhergeht, Lücken durch Automatisierung zu schließen¹⁵⁵ – ein Muster, das Kompetenzen eher abbaut als aufbaut.

Auch in Fällen wie CNET, „Gizmodo“ oder Gannett wird deutlich, dass KI-Experimente oft als Schnellschuss gestartet wurden, bevor robuste Schulungs- und Review-Strukturen etabliert waren. Dass Fehler dort zunächst nicht erkannt oder unterschätzt wurden, ist ein Indikator dafür, dass die *kollektive Urteilskraft* der Redaktionen dem Tempo der Tool-Einführung hinterherhinkte.

Fazit: Kompetenzaufbau & Training ist im Framework kein ‚Add-on‘, sondern Voraussetzung für alle anderen Felder. Ohne geteiltes Grundverständnis von KI-Arbeitsweisen bleiben Transparenz, Human-in-the-Loop oder Ethik abstrakt. Die Positivbeispiele zeigen, dass investierte Zeit in interne Experimentierräume, Trainings und Wissenstransfer sich in *robusterer Praxis* auszahlt. Die Negativmuster weisen darauf hin, dass fehlender Kompetenzaufbau zu einer gefährlichen Mischung aus Tool-Abhängigkeit und Kontrollillusion führt.

Handlungsfeld 7: Publikumsdialog & Beziehungsarbeit – Legitimität in der synthetischen Öffentlichkeit

Fokus des Frameworks: Dieses Feld verschiebt den Blick nach außen: *Wie werden Publikum, Communities und Betroffene in die Aushandlung von KI-Einsatz einbezogen?*

Resilienz meint hier die *Stabilität der Beziehung zwischen Redaktion und Öffentlichkeit* unter KI-Bedingungen.

Best Practices: *Dialogische Formate und Co-Produktion:* Die „Washington Post“ hat mit „Ask The Post AI“ ein generatives Interface eingeführt, das Nutzer:innen Fragen stellen und Antworten auf Basis der eigenen Berichterstattung geben lässt. Das Tool wird ausdrücklich als Experiment geführt; Berichte betonen, dass dessen Antworten auf kuratierter, faktengestützter Berichterstattung seit 2016 beruhen und dass das System keine Antwort ausgibt, wenn keine ausreichend relevanten Artikel gefunden wurden.¹⁵⁶ Damit wird KI als *dialogisches Zugangsformat* eingeführt, dessen Funktionsweise gegenüber Nutzer:innen offen erklärt wird.

Die Non-Profit-Redaktion „ProPublica“ steht seit Jahren für eine konsequent partizipative Praxis. Projekte wie „Free the Files“ oder langfristige Engagement-Recherchen, bei denen Leser:innen Hinweise, Dokumente und Datenspenden beisteuern zeigen, wie systematische Einbindung von Communities investigative Arbeit stärkt.¹⁵⁷ In den vergangenen Jahren hat „ProPublica“ begonnen, genau diese Engagement-Arbeit gezielt mit KI-Werkzeugen zu verknüpfen. 2025 trat die Redaktion dem vom Lenfest Institute koordinierten *AI Collaborative and Fellowship Program* bei; ein eigens eingestellter „Engineer“ soll dort erkunden, wie große Sprachmodelle helfen können, eingehende Tipps zu kategorisieren und schneller an die zuständigen Reporter:innen weiterzugeben, ohne die redaktionelle Auswahl aus der Hand zu geben.¹⁵⁸ Bereits zuvor setzte „ProPublica“ in einzelnen Recherchen *Machine Learning* und große Sprachmodelle ein – etwa zur Analyse tausender NSF-Forschungsprojekte oder zur Auswertung umfangreicher Dokumenten- und Falldaten –, wobei sämtliche KI-Ergebnisse durch Reporter:innen gegenrecherchiert wurden.¹⁵⁹ In diesem Kontext fungieren KI-gestützte Tools ausdrücklich als Verstärker von Community- und Datenrecherchen.

Auch der britische „Guardian“ nutzt seine öffentliche Kommunikation über KI-Leitlinien, Editorial Code und Ethik-Debatten, um Leser:innen einzuladen, über Erwartungen und Grenzen von KI-Einsatz mitzudiskutieren. Das stärkt nicht nur Transparenz, sondern rahmt KI als *gemeinsames Aushandlungsthema* zwischen Redaktion und Publikum.

Warnzeichen: *Reaktiver statt proaktiver Dialog:* CNET oder „Sports Illustrated“ haben gezeigt, wie Worst-Practice-Kommunikation mit dem Publikum funktioniert: nämlich überwiegend *reaktiv*. Erst nachdem externe Medien KI-Einsatz und Fehler öffentlich thematisierten, lieferten die verantwortlichen Redaktionen ihre Erklärungen und Korrekturen nach. Es gab hier kaum dialogische Formate, in denen Leser:innen Fragen stellen, Kritik einbringen oder Erwartungen an KI-Einsatz artikulieren konnten. KI blieb eine interne Infrastrukturfrage, bis eine Krise sie nach außen katapultierte.

Auch im stark automatisierten Lokaljournalismus – etwa bei Projekten, die tausende automatisch generierte Texte zu Wahlergebnissen oder Sportereignissen ausspielen – ist in der öffentlichen Kommunikation vieler Häuser bislang wenig darüber zu finden, wie diese Automatisierung gegenüber lokalen Communities erklärt wird. Damit besteht die Gefahr, dass Nutzer:innen KI-gestützte Inhalte als ‚normale‘ Berichterstattung wahrnehmen, ohne zu wissen, an wen sie sich bei Fragen oder Fehlern wenden können.

Fazit: Publikumsdialog und Beziehungsarbeit im Framework die *demokratische Dimension* von KI-Resilienz. Ist wo Redaktionen ihren KI-Einsatz offen erklären, Feedbackschleifen einbauen und partizipative Formate nutzen, wird Vertrauen nicht nur vorgetäuscht, sondern *praktisch erneuert*. Wo Kommunikation erst im Krisenfall einsetzt, untergräbt KI das bestehende Vertrauensfundament. Entscheidend ist, dass derlei Beziehungsarbeit nicht nur in ‚Innovation Labs‘ stattfindet, sondern vor allem im Alltag von Lokalteilen, Ressorts und Formaten.

Synthese: Fragmentierte Praxis und die Rolle des Frameworks

Der Praxistest entlang der sieben Handlungsfelder offenbart drei Querschnittsmuster:

1. **Fragmentierung statt Systematik:** Viele Häuser sind in einzelnen Feldern weit fortgeschritten – etwa bei Governance-Regeln („Guardian“), forensischer Verifikation („BBC Verify“) oder Kompetenzaufbau („Financial Times“, „Zetland“). Selten jedoch greifen bislang alle sieben Handlungsfelder ineinander. KI-Resilienz entsteht dann punktuell, aber nicht als *systemische Strategie* der Organisation.
2. **Governance ohne Kultur, Kultur ohne Governance:** Manche Redaktionen setzen stark auf formale Regeln und Policies, ohne diese in eine Lern- und Fehlerkultur zu übersetzen. Andere experimentieren transparent und community-orientiert, ohne klare Leitplanken und Zuständigkeiten. Das Framework zeigt, dass beides zusammenkommen muss: Regeln ohne gelebte Praxis bleiben wirkungslos, Praxis ohne Regeln bleibt fragil.
3. **Vom Einzelfall zum Muster:** Viele der beschriebenen Monita – fehlerhafte KI-Artikel, Deepfake-Risiken, Datenflüsse – werden in der öffentlichen Debatte als vermeintlich singuläre ‚Fälle‘ behandelt. Das Framework macht deutlich, dass es sich um *wiederkehrende Entscheidungssituationen* handelt: Transparenzfragen, Verantwortungszuordnung, Datennutzung, Verifikation, Ethik, Kompetenz und Dialog wiederholen sich – unabhängig vom konkreten Tool.

Gerade darin liegt der Mehrwert der KI-Resilienz-Frameworks: Sie ordnen bereits existierende Praktiken und Konflikte, machen Lücken sichtbar (zum Beispiel beim Publikumsdialog über KI oder bei der institutionellen Verankerung von Forensik) und helfen, *Insellösungen in eine kohärente Strategie zu überführen*. Die hier versammelten Best Practices zeigen, dass in jedem Handlungsfeld tragfähige Ansätze existieren. Die Warnzeichen deuten wiederum darauf hin, wo die nächsten Bruchstellen zu erwarten sind, wenn diese Felder nicht konsequent zusammengedacht werden.

Im Folgenden bauen wir darauf auf, indem wir die organisationale und gesellschaftliche Dimension noch stärker verknüpfen: *Welche Bedingungen brauchen Redaktionen, um die sieben Handlungsfelder nicht nur intern zu etablieren, sondern auch als Teil demokratischer Infrastruktur dauerhaft zu praktizieren?*

6.4 Von der lernenden Redaktion zur gesellschaftlichen Resilienz

Die eigentliche Bedeutung unseres KI-Resilienz-Frameworks entfaltet sich dort, wo journalistische Entscheidungen auf eine sensibilisierte Öffentlichkeit treffen. Für das Publikum fungiert es vor allem als Orientierungsangebot. In einer zunehmend synthetischen Medioumgebung, in der Inhalte verdaulicher, schneller und schwerer unterscheidbar werden, wächst das Bedürfnis nach Nachvollziehbarkeit. Entscheidend dabei ist: KI-Resilienz lässt sich weder hausärztlich verordnen noch unmittelbar ‚implementieren‘. Sie entsteht prozessual – durch wiederholte Auseinandersetzung, Irritation und Lernen aus Fehlern.

Der Kreislauf in Abbildung 17 operationalisiert Resilienz als *lernenden Prozess*: Zielklärung und Anpassung bewahren Standards, Anwendung nutzt Effizienz, Reflexion sichert Lernen. Genau hierin unterscheidet sie sich grundlegend von klassischen Compliance-Ansätzen. Während Regeln eine gewisse Stabilität suggerieren, operiert KI in einem Umfeld permanenter Veränderung: Modelle werden aktualisiert, Trainingsdaten erweitert, Einsatzszenarien verschieben sich. Was heute als verantwortungsvoll gilt, kann morgen schon neue blinde Flecken erzeugen.

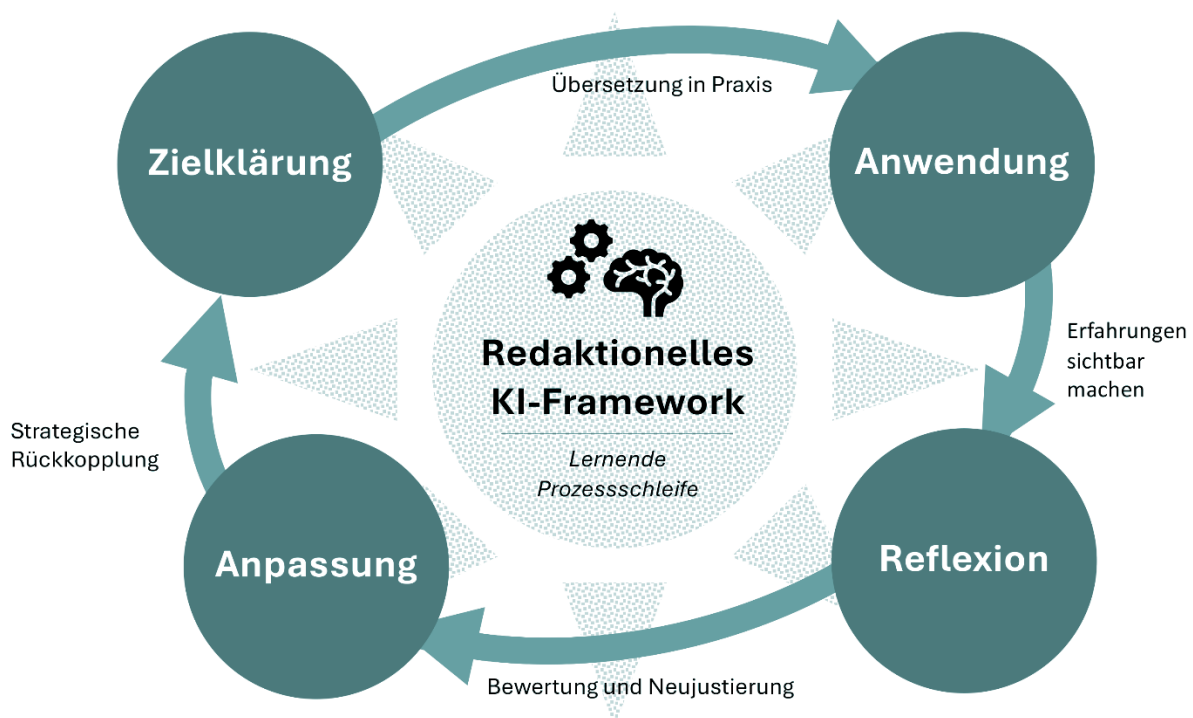


Abbildung 17: Redaktionelle KI-Governance als lernender Steuerungskreislauf (eigene Darstellung)

Lernende (resiliente) Redaktionen reagieren darauf nicht mit immer feineren Regelwerken, sondern mit der Fähigkeit zur kontinuierlichen Selbstüberprüfung. Lernende Redaktionen zeichnen sich weniger durch perfekte Voraussicht als durch institutionalisierte Reflexion aus. Sie schaffen Räume, in denen Unsicherheit thematisiert werden darf, sogar muss – jenseits von Produktionsdruck und Rechtfertigungslogiken. Feedbackschleifen, interne Reviews und kollektive Auswertungen von Verstößen sind keine Zeichen von Schwäche, sondern Voraussetzungen für Professionalität.

Gerade im Umgang mit KI entscheidet sich Resilienz dort, wo Redaktionen bereit sind, auch funktionierende Prozesse infrage zu stellen: *Warum vertrauen wir diesem System? Wo ersetzen Plausibilität und Geschwindigkeit sorgfältige Prüfung? Und wo verdrängen wir Zweifel, weil Effizienzversprechen überwiegen?* Lernende Organisationen begreifen solche Fragen nicht als Störfaktoren, sondern als Teil ihrer Routine.

KI-Resilienz bedeutet daher nicht, alle Eventualitäten vorab zu antizipieren, *sondern Strukturen zu schaffen, die Lernen ermöglichen*, wenn Unerwartetes eintritt. Redaktionen werden so selbst zu lernenden Organisationen – nicht im technischen, sondern im systemischen Sinn. Ihre Zukunftsfähigkeit liegt in der institutionellen Kompetenz, mit einer

Technologie zu arbeiten, die sich schneller verändert als jede einzelne Redaktion es je könnte – ohne dabei ihre journalistische Identität preiszugeben.

KI-Resilienz wirkt dabei weniger durch Einzelmaßnahmen als durch Konsistenz. Wenn Publika erleben, dass Transparenz nicht punktuell, sondern infrastrukturell gegeben ist; dass Verantwortung nicht im Tool einfach verschwindet, sondern beim Menschen verbleibt; dass Dialog nicht erst als Reaktion auf Kritik einsetzt, sondern eine geteilte Grundeinstellung von Medienmacher:innen ist, entsteht Beziehungskapital. Dieses Beziehungskapital ist die Voraussetzung dafür, dass Journalismus auch im KI-Zeitalter als glaubwürdige Instanz wahrgenommen wird – nicht als bloßer Content-Lieferant.

Auf gesellschaftlicher Ebene entfaltet ein solches Resilienz-Framework eine zweite, bisweilen unterschätzte Wirkung: Es wirkt medienkompetenzbildend. Redaktionen machen technologische Komplexität erfahrbar und durchschaubar(er). Indem sie ihren eigenen Umgang mit KI reflektieren und öffentlich machen, tragen sie zur Medien- und Demokratiekompetenz bei. Mediennutzende lernen nicht nur, dass KI eingesetzt wird, sondern *wie, warum – und wo ihre Grenzen liegen*. KI-Resilienz trägt somit zur kollektiven politischen Meinungs- und Willensbildung bei.

Die eigentliche Tragweite von KI-Resilienz liegt damit weit jenseits von Effizienzkategorien: Sie entscheidet darüber, ob Journalismus im KI-Zeitalter als verlässlich wahrgenommen wird – oder im Strom synthetischer Inhalte untergeht. Doch hier stößt auch das stärkste redaktionelle Framework an seine Grenzen. Vertrauen, Verantwortung und Transparenz lassen sich nicht dauerhaft auf Ebene einzelner Organisationen stabilisieren. Wenn KI zur grundlegenden Infrastruktur öffentlicher Kommunikation wird, stellen sich weitergehende globale Fragen: nach kooperativer Praxis, gemeinsamen Regeln, legitimer Machtverteilung, nach neuen Formen öffentlicher Rechenschaft. Die Frage lautet dann nicht mehr nur, wie Redaktionen KI resilient nutzen – sondern welche gesellschaftlichen Vereinbarungen notwendig sind, damit Öffentlichkeit unter KI-Bedingungen insgesamt demokratisch bleibt.

7. Ein neuer Gesellschaftsvertrag: KI-Resilienz als öffentliche Aufgabe

Worauf dieses Kapitel Antworten gibt:

1. Was aus den Analysen, Frameworks und Praxisbeispielen der Kapitel 1 bis 6 – jenseits einzelner Anwendungsfälle oder Organisationsfragen folgt.
2. Warum KI-Resilienz keine interne Management- oder Redaktionsaufgabe ist, sondern zu einer öffentlichen und demokratischen Verantwortung wird.
3. Welche normativen Setzungen und strategischen Konsequenzen sich daraus für Journalismus, Medienpolitik, Ausbildung und Öffentlichkeit ergeben.
4. Wie ein neuer Gesellschaftsvertrag im KI-Zeitalter gedacht werden kann, ohne technologische Heilsversprechen oder Kontrollillusionen zu reproduzieren.

7.1 Warum ein neuer Gesellschaftsvertrag notwendig ist

Die Entscheidung bei der „Washington Post“ war gefallen, bevor viele in der Redaktion sie im Februar 2026 zur Kenntnis nehmen konnten. In Interviews erklärte sich Chefredakteur Matt Murray dafür verantwortlich und nannte sie Teil eines Plans, den er gemeinsam mit seinem Führungsteam ausgearbeitet habe: Die Auslandsberichterstattung werde reduziert, sagte er – Reporter sollten zwar in fast einem Dutzend Standorten erhalten bleiben, doch ganze Regionen würden wegfallen. Journalisten und Redakteure im Nahen Osten, in Indien und Australien verloren ihre Stellen.

Wie die *New York Times* rekonstruierte, lösten einige dieser Schritte im Newsroom Fassungslosigkeit aus.¹⁶⁰ Ein Ukraine-Korrespondent wurde entlassen, noch während er aus dem Kriegsgebiet berichtete; die gesamte Fotoabteilung der Zeitung wurde gestrichen; ein Sportreporter, der für die Olympischen Winterspiele in Italien vor Ort war und abgezogen werden sollte, kündigte an, trotz Entlassung weiter Texte zu liefern. Selbst innerhalb der Führungsriege regte sich Widerstand. Peter Finn, Leiter der Auslandsredaktion, bat darum, selbst entlassen zu werden, statt am Restrukturierungsplan für die Streichung von 300 journalistischen Stellen mitzuwirken, als ihm deren tatsächlicher Umfang klar wurde.

Der drastische Stellenabbau bei der „Post“ vollzieht sich damit nicht als abstrakte Restrukturierung, sondern als Abfolge konkreter Einzelentscheidungen – getroffen von Menschen, dekretiert per E-Mail und bis in entlegene Kriegsgebiete und Nachrichtenlagen hinein spürbar. Er markiert einen harten Einschnitt, der sich nicht allein aus ökonomischem Druck erklären lässt. Es ist ein Moment, der eine strategische Neuvermessung dessen

sichtbar macht, was die „Post“ als einstiges *Paper of Record* künftig noch zu leisten vermag – und was nicht mehr.

Dieser Einschnitt ist kein isolierter Managementfehler, sondern Symptom eines tieferliegenden Veränderungsprozesses: Der Zugang zu Öffentlichkeit wird neu organisiert – und mit ihm die Frage, wer unter KI-Bedingungen noch Verantwortung trägt und wofür. In dieser Logik verlieren ausgerechnet jene journalistischen Leistungen systematisch an Wert, die Zeit, Kontext und Präsenz erfordern: Auslandsberichterstattung, Recherche vor Ort, visuelle Dokumentation. Die Sparwelle bei der „Post“ fällt damit nicht zufällig, sondern strukturell folgerichtig in eine Phase, in der journalistische Produktionsprozesse nicht nur von KI punktuell unterstützt werden, sondern der Zugang zu Öffentlichkeit sich generell neu sortiert.

Zusammengenommen ergibt sich ein Befund, der über diesen Kahlschlag als Einzelepisode hinausweist: KI wirkt vor allem als publizistische Selektionsmacht, die Effizienz, Sichtbarkeit und Relevanz neu verteilt – und dabei gerade jene journalistischen Leistungen entwertet, die schwer skalierbar sind. Gewiss ist, dass KI nicht erst kommt, sondern wie bis hierher gezeigt längst da ist: KI-basierte Suche, Aggregation und Zusammenfassung verändern auch hierzulande immer mehr die Mediennutzung und vor allem den Zugang zu Nachrichten, entkoppeln Inhalte vollständig von ihren Absendern und verschieben Aufmerksamkeit weg von journalistischen Marken hin zu *KI-Interfaces*.

Doch genau hierin liegt der zentrale blinde Fleck: In vielen Redaktionen wird KI vor allem als Werkzeug behandelt – als Tool für mehr Produktivitätsgewinn, möglicherweise als Sparinstrument oder als Innovations-Statussymbol. Damit wird KI zur Verantwortungsfrage. Nicht mehr entscheidend ist, welche Tools erlaubt sind, sondern ob demokratische Öffentlichkeit unter KI-Bedingungen überhaupt noch verantwortlich organisiert ist – und welche Rolle der Journalismus dabei spielen soll. Wo diese Bedingungen nicht mehr erfüllt sind, verliert Öffentlichkeit ihre Fähigkeit zur Unterscheidung – und damit ihre demokratierelevante Funktion.

Die vorangegangenen Kapitel haben gezeigt: KI-Resilienz entsteht dort, wo Verantwortung nicht behauptet, sondern organisatorisch getragen wird – durch Routinen, Zuständigkeiten und Lernformate. Unter realen Bedingungen jedoch – Zeitdruck, Ressourcenknappheit, Unsicherheit – gelingt genau das nicht automatisch, sondern wird zur offenen Gestaltungsfrage. Vor diesem Hintergrund wird deutlich, warum KI nicht länger nur als Produktivitätsmaschine oder als Totalbedrohung verstanden werden darf, sondern als Herausforderung für Medienvertrauen, demokratisches Miteinander und professionelle Handlungsfähigkeit.

Im Zentrum dieses Kapitels steht daher abschließend eine einfache, aber folgenreiche These: *KI-Nutzung ist nur dort legitim, wo Verantwortung nicht verpufft, sondern organisiert und transparent bleibt.*

Denn genau daran entscheidet sich, ob Journalismus im KI-Zeitalter eine tragende Institution öffentlicher Orientierung bleibt – oder schleichend ausgehöhlt wird. Verantwortung wird allzu häufig ausgelagert – an technische Systeme, organisationale Sachzwänge oder ökonomische Logiken, die selbst keiner öffentlichen Rechenschaft unterliegen. Ein KI-Resilienz-Framework setzt deshalb einen neuen Gesellschaftsvertrag voraus: einen Rahmen, der klärt, ob und wie journalistische Verantwortung in einer KI-vermittelten Öffentlichkeit noch verbindlich zugerechnet werden kann.

Dementsprechend ist KI auch nicht einfach eine weitere Medientechnologie, sie ist eine *neue Infrastruktur öffentlicher Ordnung*. Sie wirkt dort, wo Entscheidungen vorbereitet, priorisiert und plausibilisiert werden – meist ohne kenntlich zu werden. Diese Unsichtbarkeit ist kein technisches Detail, sondern eine demokratische Herausforderung. Wo Entscheidungen nicht mehr nachvollziehbar sind, verliert Öffentlichkeit ihre Orientierungsfunktion. Wo Verantwortung hinter Systemen verschwindet, wird Vertrauen nicht aufgebaut, sondern schrittweise abgebaut.

Im Journalismus manifestiert sich diese Verschiebung besonders deutlich. Er ist nicht nur Nutzer von KI, sondern ein zentraler Ort, an dem sich entscheidet, ob öffentliche Kommunikation wahrhaftig und in sich konsistent bleibt.

7.2 Der normative Kern:

Resilienz als Lernfähigkeit, nicht als Robustheit

Der neue Gesellschaftsvertrag knüpft Resilienz an Lernfähigkeit – auf individueller, organisationaler und gesellschaftlicher Ebene. Journalist:innen, ihre Redaktionen, die Medienbranche und die Gesellschaft insgesamt müssen lernen, unter Unsicherheit begründete Entscheidungen zu treffen und ihre eigene Verantwortung sichtbar zu halten.

Was heißt das konkret?

Wer KI verantwortungsvoll nutzen will, muss sich von einer Hoffnung verabschieden: dass sich Risiken vollständig vermeiden lassen. Fehler sind unter KI-Bedingungen (noch häufig) der Normalfall. Entscheidend ist dabei weniger, *ob* sie auftreten – sondern *wie* mit ihnen umgegangen wird. Resilienz setzt voraus, dass Einzelne urteilen können, Organisationen Verantwortung absichern und die Gesellschaft neue Rahmenbedingungen schafft,

in denen Lernen aus eigenen Fehlern möglich ist, ohne Sanktion oder Tabuisierung – und ohne dass eine individuelle KI-Nutzung sofort als persönliches Risiko gelesen wird.

Hieraus speist das vorgeschlagene KI-Resilienz-Framework seine normative Relevanz: Es fungiert als Entscheidungsarchitektur für Situationen, in denen Redaktionen unter Zeitdruck, zwischen widersprüchlichen Zielen und ohne klare Antworten handeln müssen – und dennoch niemand mit der Verantwortung alleine gelassen bleiben sollte. Das Framework macht Verantwortung institutionell implementierbar. Es schafft Räume, in denen Entscheidungen begründet, überprüft und – wenn nötig – korrigiert werden können.

Diese Praxis kann nicht hinter Redaktionstüren verborgen bleiben. In einer fragmentierten Öffentlichkeit vertrauen Menschen nicht mehr automatisch Institutionen, nur weil diese einen althergebrachten Namen haben. Vertrauen wächst, wenn Entscheidungen erklärt werden, bevor danach gefragt wird – proaktiv, nicht defensiv.

Ein Beispiel: Wenn eine Redaktion KI nutzt, um Lokalnachrichten aus strukturierten Daten zu generieren – etwa Wahlergebnisse oder Wetterdaten –, reicht es nicht zu sagen ‚dieser Text ist KI-generiert‘. Öffentlich verantwortbar wird der Einsatz erst, wenn erklärt wird: *Welche Daten wurden verwendet? Wer hat sie geprüft? Wer trägt die Verantwortung, wenn etwas schief geht? Und wo liegen die Grenzen dieses Verfahrens?* Nur so können Mediennutzer:innen einordnen, ob und wann sie dem Ergebnis vertrauen – oder skeptisch nachhaken sollten.

Ähnlich gilt: Eine Redaktion nutzt KI, um Interviewtranskripte zusammenzufassen. Öffentlich verantwortbar wird das nur, wenn transparent wird, dass die Zusammenfassung geprüft wurde, dass die Originalquelle zugänglich bleibt und dass niemand sich darauf verlässt, die KI habe ‚schon alles richtig‘ erfasst. Verantwortbarkeit bedeutet: Wer Inhalte wahrnimmt, kann verstehen, wie diese Information entstanden ist – und wer dafür einsteht.

Resilienz wird so automatisch zu einer demokratischen Kategorie: Sie beschreibt nicht, wie sich Journalismus den Wandel durch KI umarmen oder sich dagegen wehren kann. Sie beschreibt, wie diese Transformation im Digitalen so gestaltet wird, dass Öffentlichkeit in der Lage ist, diese nachzuvollziehen, zu bewerten und zu kritisieren, statt sich ihr ohnmächtig ausgeliefert zu fühlen.

7.3 Pflichten des Journalismus: Neue professionelle Zumutungen

Ein Gesellschaftsvertrag ist kein Appell, sondern – im besten Sinne – eine *professionelle Zumutung*. Die erste Zumutung betrifft die Frage, worum es überhaupt geht: KI-Resilienz ist keine Tool-, sondern eine Steuerungs- und Governance-Frage. Solange Redaktionen KI-Einsatz als Problem einzelner Anwendungsfälle behandeln, bleibt Verantwortung zersplittert. Weder Qualitätsgewinne noch Fehlleistungen lassen sich sinnvoll auf konkrete Systeme zurückführen. Worauf es ankommt, ist vielmehr eine strukturelle Frage: *Können Organisationen Entscheidungsmacht, Kontrolle und Verantwortung unter KI-Bedingungen dauerhaft organisieren?*

Wie unterschiedlich Redaktionen diese Frage beantworten – von regel- und governance-orientierten Mustern über kompetenz- und kulturgetriebene Ansätze bis hin zu primär ökonomischen Logiken –, zeigt bereits die alltägliche Medienpraxis. Damit verschiebt sich auch, was Transparenz bedeutet. Offenlegung allein schafft Sichtbarkeit, aber kein Vertrauen. Kennzeichnungen, Ethik-Statements oder Transparenzseiten wirken nur, wenn sie mit klaren Zuständigkeiten, überprüfbaren Entscheidungswegen und einer erkennbaren Fehlerkultur verbunden sind. Transparenz ist keine Dokumentationsleistung, sondern eine Form professioneller Beziehungsarbeit.

Unter KI-Bedingungen wird deutlich, welche Ressource in Redaktionen knapp wird: Urteilskraft. Inhalte lassen sich skalieren, Entscheidungen aber nicht. KI-Resilienz bedeutet also, diese Ressource organisational zu schützen – gegen Zeitdruck, gegen Automatisierungslogiken und gegen die Verführung plausibler Ergebnisse. Nicht Geschwindigkeit wird zur Währung journalistischer Qualität, sondern die Fähigkeit, Entscheidungen nachvollziehbar und überzeugend zu begründen.

Diese Anforderungen lassen sich nicht in abschließende Regelwerke pressen. Modelle, Trainingsdaten und Einsatzkontexte verändern sich kontinuierlich. KI-Resilienz entsteht daher prozessual: durch Reviews, Reflexionsräume und die Bereitschaft, auch funktionierende Routinen infrage zu stellen. Entscheidend ist nicht Regelkonformität, sondern die Fähigkeit zur fortlaufenden Selbstkorrektur. Redaktionelle Verantwortung verbleibt daher auch nicht in den Redaktionsstuben. Journalismus wendet KI nicht nur an, er wirkt als *öffentlicher Referenzraum*. Wie Redaktionen KI nutzen, erklären und begrenzen, prägt gesellschaftliche Erwartungen an Technologie insgesamt, erst recht beim Hype-Thema KI. Redaktionelle Entscheidungen setzen – ob bewusst oder unbewusst – gesellschaftliche Normen.

In einer KI-vermittelten Öffentlichkeit hängt demokratische Resilienz darüber hinaus daran, dass Informationen noch unterscheidbar bleiben. Wo sich Entstehungsprozesse

nicht mehr nachvollziehen lassen, erodiert das allgemeine Vertrauen in ‚die Medien‘. KI-Resilienz betrifft damit nicht nur den Journalismus, sondern die Fähigkeit von Öffentlichkeit, sich intellektuell zu behaupten.

Ein neuer Gesellschaftsvertrag für KI beginnt nicht bei den Verboten. Am Anfang steht Verantwortung. Nicht Kontrolle allein entscheidet über die gesellschaftliche Wirkung von KI, sondern wie bewusst deren Zuständigkeiten verortet werden. Diese Erkenntnisse markieren keinen Verlust journalistischer Autonomie. Im Gegenteil: Sie schaffen erst die Voraussetzungen, unter denen sich Autonomie unter KI-Bedingungen erneuern kann.

7.4 Konsequenzen und Ableitungen nach Verbundebenen

Ein Gesellschaftsvertrag für das KI-Zeitalter bindet alle ein, die an öffentlicher Kommunikation beteiligt sind. Er adressiert die strukturellen Bedingungen, unter denen journalistische Verantwortung überhaupt wahrgenommen werden kann. Wenn KI zur Infrastruktur öffentlicher Kommunikation wird, müssen mehrere institutionelle Ebenen zusammenspielen. KI-Resilienz entsteht demnach im Verbund – dort, wo redaktionelle Praxis, medienpolitische Rahmenbedingungen, Ausbildungssysteme und Öffentlichkeit aufeinander abgestimmt sind. Fehlt eine dieser Ebenen, wird der Verantwortungs-Nimbus entweder überspannt, behauptet oder nur symbolisch aggregiert. Die folgenden Ableitungen unterscheiden vier miteinander verbundene Ebenen: Redaktion/Organisationen, Regulierung, Qualifikation und öffentliche Orientierung.

Redaktion & Medienorganisationen

Für Redaktionen und Medienhäuser stellt die KI-Nutzung vor allem eine Führungs- und Organisationsfrage dar. Verantwortung muss institutionell verankert werden, statt sie situativ zu verteilen. Der Hebel liegt in der Verschiebung von individueller Nutzung zu organisierter Praxis. Frameworks machen redaktionelle Autonomie belastbar, indem sie wiederkehrende Entscheidungssituationen strukturieren und Zuständigkeiten klären.

Wirksame *Human-in-the-Loop*-Strukturen gewinnen hier an Gewicht, sind aber – wie gezeigt – nicht der Weisheit letzter Schluss: Die abschließende Verantwortung menschlicher Redaktionsmitglieder muss auch unter Zeitdruck greifen. Das setzt klare Zuständigkeiten voraus und eine Organisationskultur, in der Zweifel, Rückfragen und Korrekturen als professionelle Standards gelten. Hier liegt eine zentrale Sollbruchstelle vieler Redaktionen: Unter hoher Arbeitsverdichtung wird Prüfung zur Individualleistung und Zweifel gilt

als Zeitverlust. Resilienz verlangt deshalb Schutzräume und Routinen, die Rückfragen auch dann ermöglichen, wenn die journalistische Produktion unter Wind steht.

Transparenz wird in diesem Kontext zur dauerhaften Organisationsleistung. Sie muss als verlässliche Praxis erkennbar sein. Sichtbare Verantwortung, erklärbare Entscheidungen und nachvollziehbare Korrekturen stärken das Vertrauen des Publikums und die interne Orientierung gleichermaßen. Resiliente Redaktionen lernen aus Grenzfällen und Versäumnissen – Beispiel ZDF, „Spiegel Online“ oder Ippen – idealerweise systematisch – und im besten Fall mit ihnen die gesamte Branche. Institutionalisierte Reflexion – etwa in Reviews, Fallanalysen oder kollektiven Lernformaten – entlastet Individuen und stärkt Organisationen. Ohne solche Strukturen bleibt Verantwortung unter KI-Bedingungen eine Leerstelle.

Die leitende Frage dieser Ebene lautet: *Wie organisieren Redaktionen Entscheidungsmacht, Zuständigkeiten und Verantwortung in Bezug auf KI so, dass journalistische Urteilskraft auch unter Zeit- und Effizienzdruck erhalten bleibt?*

Medienpolitik, Regulierung und Förderung

Mit KI-Resilienz verschiebt sich der medienpolitische Fokus grundlegend. Medienpolitik muss Bedingungen schaffen, unter denen verantwortliche Nutzung überhaupt möglich wird. Regulierung, die sich auf die Kontrolle konkreter Tools beschränkt, läuft der technologischen Entwicklung zwangsläufig hinterher. Ein Perspektivwechsel wird nötig: weg von reaktiver Aufsicht, hin zu struktureller Ermöglichung. Denn Medienpolitik beeinflusst, ob Organisationen über Zeit, Ressourcen und institutionellen Rückhalt verfügen, um Verantwortung tatsächlich wahrzunehmen.

Regulierung setzt Grenzen, wo demokratische Mindeststandards berührt sind. Förderung entscheidet darüber, ob diese Standards im Alltag lebbar sind. Eine Politik, die Transparenz, Sorgfalt und Lernfähigkeit fordert, diese aber nicht finanziell und organisatorisch absichert, erzeugt lediglich symbolische Verantwortung ohne praktische Wirkung. Förderpolitik spielt deshalb eine Schlüsselrolle für KI-Resilienz. Bislang fließen Mittel häufig in technische Innovation und Start-ups, Produktentwicklung oder kurzfristige Skalierung. Resilienz erfordert völlig andere Prioritäten: Investitionen in Organisation, Weiterbildung, gemeinsame Infrastruktur und reflexive Praxis. Gefördert werden sollte, was Verantwortung dauerhaft tragfähig macht.

Regulierung gewinnt an Wirksamkeit, wenn sie Resilienz als Schutzprinzip anerkennt. Journalistische KI-Governance kommt dabei eine demokratische Funktion zu –

vergleichbar mit Quellenschutz oder redaktioneller Unabhängigkeit. Wo Politik diesen Zusammenhang erkennt, kann sie Journalismus gezielt stabilisieren.

Die leitende Frage dieser Ebene lautet: *Unter welchen Bedingungen lässt sich KI verantwortlich steuern, fördern und regulieren?*

Journalistische Ausbildung und Weiterbildung

In Ausbildung und Weiterbildung wäre ein rein technischer Zugriff verantwortungslos. KI-Kompetenz als Toolwissen oder Systemverständnis zu vermitteln bleibt folgenlos, wenn sie nicht mit Urteilkraft, Verantwortung und organisationalem Denken verbunden wird. Unter KI-Bedingungen reicht es nicht mehr aus zu wissen, wie KI-Systeme funktionieren. Entscheidend wird, wer ihre Ergebnisse verantwortet – und auf welcher Grundlage.

Das hat Konsequenzen für Ausbildungslogiken. Punktuelle Schulungen oder isolierte Tool-Trainings erzeugen Anpassungsfähigkeit. Resilienz entsteht mithin durch kontinuierliche Lernprozesse, in denen Entscheidungskonflikte, Grenzfälle und Unsicherheiten systematisch reflektiert werden. KI-Resilienz entsteht vor allem dort, wo technische, ethische und organisationale Fragen zusammengedacht werden. Weiterbildung sollte professionelle Urteilkraft daher in Grenzfällen trainieren: *Was gilt als hinreichend geprüft, wann wird KI-Nutzung riskant, und wer entscheidet im Konfliktfall?*

Besonders problematisch ist die Individualisierung von KI-Kompetenz. Wenn Herrschaftswissen über Systeme bei wenigen Spezialist:innen konzentriert ist, entstehen neue Abhängigkeiten und informelle Machtstrukturen. Verantwortung wird deloziiert, statt geklärt. Ausbildung muss dem ein kollektives Kompetenzverständnis entgegensetzen: gemeinsames Lernen, geteilte Entscheidungslogiken, transparente Zuständigkeiten. Weiterbildung wird so selbst zur infrastrukturellen Voraussetzung von KI-Resilienz. Sie stärkt die Fähigkeit, mit dauerhafter Unsicherheit umzugehen. Diese Fähigkeit entscheidet darüber, ob journalistische Professionalität unter beschleunigten Bedingungen erhalten bleibt.

Die Leitfrage dieser Ebene lautet: *Wie bilden wir Journalist:innen für eine digitale Praxis aus, in der sich Werkzeuge permanent verändern, publizistische Verantwortung aber bestehen bleiben muss?*

Öffentlichkeit, Publikum und demokratische Kultur

Die Belastbarkeit von KI-Resilienz-Frameworks im Journalismus entscheidet sich im Verhältnis zur Öffentlichkeit. Dort zeigt sich, ob Redaktionen auch unter KI-Bedingungen als glaubwürdige Instanzen öffentlicher Orientierung wahrgenommen werden – oder ob Vertrauen durch synthetisch erzeugte Plausibilität ausgehebelt wird. In einer Umgebung, in der Inhalte immer schneller und glatter werden und sich kaum noch unterscheiden lassen, wächst das Bedürfnis nach Einordnung. Diese Einordnung entsteht dort, wo Verantwortung in aller Deutlichkeit sichtbar übernommen wird.

Hier ist der Journalismus in seiner zentralen Funktion angesprochen: Er macht technologische Vermittlung sichtbar, indem er seine eigenen Praktiken offenlegt. Er erklärt die Welt – und zugleich, unter welchen Bedingungen die vermittelten Weltbilder zustande kommen. Journalismus wird damit zum Übersetzer technologischer Macht – durch konsequente Transparenz. Diese Form der Beziehungsarbeit ist anspruchsvoll: Sie verlangt, Entscheidungen zu begründen, Grenzen zu benennen und Ambiguität auszuhalten. Für das Publikum ist das ebenfalls eine Zumutung: Vertrauen entsteht aus der Bereitschaft, sich bewusst mit redaktionellen Entscheidungen auseinanderzusetzen, um sie nachzuvollziehen zu können.

Gerade im lokalen oder serviceorientierten Journalismus wird diese Beziehungspflege gerade jetzt wichtiger als jemals zuvor. Dort entsteht Glaubwürdigkeit durch Wiedererkennbarkeit, Dialog und Verlässlichkeit. KI-Resilienz bedeutet hier, KI-gestützte Automatisierung zu erklären – oder bewusst zu begrenzen, wenn sie diese wertvollen Beziehungen gefährdet. Wo sich Entstehungsprozesse von Information und Verantwortlichkeiten nicht mehr direkt nachvollziehen lassen, wird Kritik wirkungslos und Vertrauen beliebig. KI-Resilienz ist damit eine Voraussetzung demokratischer Kultur.

Die leitende Frage dieser Ebene lautet: *Wie bleibt Öffentlichkeit urteilsfähig, wenn Information zunehmend KI-gestützt vermittelt wird?*

7.5 Die Rolle des Staates: Ermöglicherin statt Technokratin

Ein neuer Gesellschaftsvertrag für das KI-Zeitalter verlangt, auch die staatliche Verantwortung neu zu lokalisieren – als eigenständige demokratische Funktion, ohne medienpolitische Regulierungsfantasien zu bedienen. Der Staat trägt Mitverantwortung für die Bedingungen öffentlicher Kommunikation und dafür, ob und wie Verantwortung unter KI-Einflüssen tatsächlich wahrgenommen werden kann.

In der aktuellen KI-Debatte wird diese Verantwortung häufig auf Regulierung verkürzt: Gesetze, Aufsicht, Compliance. Diese Instrumente sind notwendig, greifen aber oftmals ins Leere. Sie setzen an sichtbaren Anwendungen, Risiken oder Verstößen an – die strukturellen Voraussetzungen demokratischer Handlungsfähigkeit aber bleiben außen vor. KI-Resilienz verlangt mehr als Kontrolle. Sie verlangt staatliche Intervention in institutionelle Lernfähigkeit: Weiterbildung, Kompetenzaufbau und den Schutz professioneller Entscheidungsräume.

Wenn redaktionelle Organisationen unter permanentem ökonomischem, zeitlichem und technologischem Druck stehen, wird Verantwortung leicht zur Überforderung. Ein resilienzorientierter Staat erkennt daher journalistische Urteilskraft als öffentliches Gut an – vergleichbar mit der Pressefreiheit oder redaktioneller Unabhängigkeit – und schafft Rahmenbedingungen, die ihre Ausübung auch dauerhaft ermöglichen. Ebenso zentral ist der Aufbau gemeinsamer Infrastrukturen für kollektives Lernen. Einzelne Redaktionen können die langfristigen Wirkungen von KI-Systemen nicht allein stemmen. Notwendig sind Wissensräume, Forschungsverbünde, Reallabore, Resilienznetzwerke und Austauschformate, in denen Erfahrungen systematisch gebündelt, reflektiert und zurückgespiegelt werden. KI-Resilienz entsteht im gemeinsamen Umgang mit Unsicherheit.

Gleichzeitig muss der Staat der Versuchung *technokratischer Übersteuerung* widerstehen. Je komplexer KI-Systeme werden, desto größer wird der Wunsch nach scheinbar pragmatischen Lösungen: automatisierte Prüfungen, standardisierte Risiko-Scores, formalisierte Transparenzpflichten. Solche Instrumente können unterstützen – sie dürfen jedoch nicht die Illusion verstärken, Verantwortung lasse sich delegieren. Ein resilienzorientierter Staat akzeptiert daher, dass sich manches seiner Steuerung entzieht. Er setzt auf Verfahren statt auf Versprechen, auf Lernprozesse statt auf Endlösungen. Darin unterscheidet sich ein demokratischer Gesellschaftsvertrag grundlegend von autoritären Kontrollmodellen.

7.6 Macht, Asymmetrie und öffentliche Orientierung

Ein blinder Fleck von KI-Debatten ist in der Regel die Machtfrage. Wer über Rechenleistung, Daten, Modelle und Deutungsrahmen verfügt, prägt öffentliche Wirklichkeit, lange bevor formale Regulierung greift (vgl. Kap. 5). Diese Macht wirkt strukturell: Sie entscheidet darüber, was sichtbar wird, was plausibel erscheint und welche Optionen letztlich denkbar sind. Diese Asymmetrie hat allerdings konkrete wirtschaftliche Folgen: Wenn KI-Oberflächen Antworten direkt liefern, sehen Menschen journalistische Angebote oft gar

nicht mehr. Das schwächt Reichweite, Einnahmen und damit die Eigenständigkeit des Journalismus – Stichwort: ‚People Zero‘.

Für demokratische Öffentlichkeit liegt das Problem daher tiefer. Wo technologische Vermittlung unsichtbar wird, verliert Öffentlichkeit ihre Fähigkeit zur digitalen Selbstermächtigung. Kritik setzt Kenntnis voraus, Widerspruch braucht nachvollziehbare Bezugspunkte. Verschwinden diese, wird Macht vorausgesetzt statt kontrolliert. KI-Resilienz verschiebt diesen Fokus von technischer Leistungsfähigkeit auf institutionelle Verantwortung. Sie fragt weniger, *wie* leistungsfähig Systeme funktionieren. Sie fragt vielmehr, *wer* für ihre Wirkungen verantwortlich ist – organisatorisch, professionell und öffentlich. Journalismus steht dabei in einer doppelten Rolle: Er ist Teil jener Machtkonstellation, weil er KI nutzt, integriert und – auf seine Weise – ‚normalisiert‘. Zugleich bietet er einen der wenigen Orte, an denen diese Macht überhaupt noch sichtbar gemacht werden kann. Diese Selbstbezüglichkeit wird zur demokratischen Ressource, sofern sie reflektiert und offengelegt wird.

Ein neuer Gesellschaftsvertrag für KI bedeutet, Öffentlichkeit als Unterscheidungsraum zu schützen. Wo technische Plausibilität an die Stelle öffentlicher Begründung tritt, wird öffentliche Kritik entwertet. KI-Resilienz setzt dem eine andere Logik entgegen: Erklärbarkeit wiegt mehr als Effizienz, Verantwortlichkeit mehr als Skalierung, Beziehung mehr als Autorität. Öffentliche Orientierung entsteht somit durch nachvollziehbare Entscheidungen, perfekte Systeme sind dafür nicht nötig. An diesem generellen Punkt entscheidet sich, ob KI demokratisch eingebettet oder autoritär wirksam wird.

7.7 Schlussgedanke: KI-Resilienz als demokratische Infrastruktur

KI-Resilienz wehrt nichts ab. Sie gestaltet. Sie akzeptiert technologische Entwicklung, überlässt ihr aber die Richtung nicht. Sie nimmt Unsicherheit auf, mystifiziert sie jedoch nicht. Der neue Gesellschaftsvertrag im KI-Zeitalter regelt insofern kein Verhältnis zwischen Mensch und Maschine. Er beschreibt, wie Verantwortung neu organisiert wird, wenn Technik plötzlich Entscheidungen vorbereitet, diese aber nicht legitimieren kann. Er bindet Redaktionen, Öffentlichkeit und Staat an die gemeinsame Aufgabe, Orientierung unter Bedingungen technischer Vermittlung zu sichern.

In diesem Verständnis wird KI-Resilienz zur demokratischen Infrastruktur. Dort, wo Entscheidungen transparent bleiben, Zuständigkeiten sichtbar sind und Lernen institutionell verankert wird, bleibt Öffentlichkeit fähig zu unterscheiden und zu urteilen – auch dann, wenn technische Systeme immer mehr an Bedeutung gewinnen. Journalismus kann in diesem Gefüge eine zentrale Rolle übernehmen. Er wirkt als institutionelles Korrektiv, die

Rolle des technischen Vorreiters muss er nicht annehmen. Er zeigt, dass sich Macht begrenzen lässt, wenn Entscheidungen begründet, Prozesse offengelegt und Zuständigkeiten nicht delegiert werden. Damit schließt sich der Kreis zur Ausgangsdiagnose: Resilienz entsteht durch überprüfbare Praxis, die vor allem unter Unsicherheit funktioniert. Eine weitere KI-Leitlinie genügt dafür nicht.

Ein solcher Gesellschaftsvertrag beginnt bei Verantwortung. Er entsteht dort, wo Institutionen ihre Entscheidungsfähigkeit sichern, wo Öffentlichkeit nachvollziehen kann, wie Informationen zustande kommen und demokratische Selbststeuerung unter KI-Bedingungen möglich bleibt. Dieses Whitepaper formuliert dahingehend eine eindeutige Position: Als *Bedingung demokratischer Zukunftsfähigkeit* lässt sich KI-Resilienz nur gemeinsam gestalten.

Endnoten

- ¹ Vgl. Uth, Bernadette (2025): What does the audience expect when interacting with news? A Q-sort Study. In: Journalism Practice, <https://doi.org/10.1080/17512786.2025.2551986>; America Amplified (2025): Community Engagement Journalism Playbook 2025. America Amplified, <https://americaamplified.org/wp-content/uploads/2025/04/America-Amplified-Community-Engagement-Journalism-Playbook-2025.pdf>; Radcliffe, Damian (2024): Advancing community-centered journalism: A practical guide for a more equitable and service-oriented form of journalism. University of Oregon: School of Journalism and Communication, https://www.researchgate.net/publication/385726191_Advancing_Community-Centered_Journalism; Karolidou, Eva/Matsiola, Maria (2025): The role of online communities in transforming news production: challenges and opportunities. Online Information Review, 50(1), <https://doi.org/10.1108/OIR-02-2025-0064>
- ² Vgl. Banfi, Jaime Abello u.a. (2025): Bricks and mortar: the new wave of houses of journalism. International Journalism Festival, 12.04.2025, <https://www.journalismfestival.com/programme/2025/bricks-and-mortar-the-new-wave-of-houses-of-journalism>
- ³ Thompson, Stuart A./Hsu, Tiffany/Myers, Steven Lee (2026): Chaos in Minneapolis Exposes an Internet at War With Truth. In: The New York Times, 03.02.2026, <https://www.nytimes.com/2026/02/03/technology/minneapolis-ai-disinformation-misinformation-truth.html>
- ⁴ Generative KI-Inhalte wie Deepfakes erzeugen nicht nur technische Herausforderungen, sondern untergraben das Vertrauen in traditionelle Medien bzw. verstärken bei Nutzer:innen das Gefühl, die Grenze zwischen Realität und Fiktion verliere ihre Verlässlichkeit (vgl. Hynek, Nik/Gavurova, Beata/Kubak, Matus (2025): Risks and benefits of artificial intelligence deepfakes: Systematic review and comparison of public attitudes in seven European Countries. In: Journal of Innovation & Knowledge, Vol. 10 (5), <https://doi.org/10.1016/j.jik.2025.100782>; López-Borrull, Alexander/Lopezosa, Carlos (2025): Mapping the Impact of Generative AI on Disinformation: Insights from a Scoping Review. In: Publications 2025, 13(3), 33, <https://doi.org/10.3390/publications13030033>; Peña-Alonso, Urko/Peña-Fernández, Simón/Meso-Ayerdi, Koldobika (2025): Journalists' Perceptions of Artificial Intelligence and Disinformation Risks. In: Journal and Media 2025, 6(3), <https://doi.org/10.3390/journalmedia6030133>; Park, Seyeon/Nan, Xialoli (2025): Generative AI and misinformation: A scoping review of the role of AI in the generation, detection, mitigation, and impact of misinformation. In: AI & Society, <https://doi.org/10.1007/s00146-025-02620-3>

- ⁵ Auf die Frage nach einem Szenario, wenn „die Systeme böartig werden“, antwortet Geoffrey Hinton im Interview mit dem US-TV-Magazin „60 Minutes“: „They will be able to manipulate people, right? And these will be very good at convincing people ‘cause they’ll have learned from all the novels that were ever written, all the books by Machiavelli, all the political connivances, they’ll know all that stuff. They’ll know how to do it. [...] One of the ways in which these systems might escape control is by writing their own computer code to modify themselves. And that’s something we need to seriously worry about.“ (zitiert nach Pelley, Scott (2024): “Godfather of Artificial Intelligence” Geoffrey Hinton on the promise, risks of advanced AI. In: CBS 60 Minutes, 16.06.2024, <https://www.cbsnews.com/news/geoffrey-hinton-ai-dangers-60-minutes-transcript/>)
- ⁶ Vgl. Weichert, Stephan (2025a): Auf in die Schlacht gegen Dark Tech. In: taz – die tageszeitung, Medien, S. 18; Weichert, Stephan (2025b): Es gibt keine künstliche Intelligenz. In: taz – die tageszeitung, S. 15; sowie beispielhaft Andree, Martin (2023, Neuaufl. 2025): Big Tech muss weg! Die digitalen Monopole zerstören Demokratie und Wirtschaft – wir werden sie stoppen. Frankfurt/Main: Campus Verlag; Zuboff, Shoshana (2019): The age of surveillance capitalism: The fight for a human future at the new frontier of power. New York, NY: PublicAffairs; Suleyman, Mustafa/Bhaskar, Michael (2023): The coming wave. Technology, power, and the twenty-first century’s greatest dilemma. New York, NY: Crown; Morozov, Evgeny (2024): Freedom as a service: The new digital serfdom. New York, NY: Farrar, Straus and Giroux; Andree, Martin (2025): Krieg der Medien: Dark Tech und Populisten übernehmen die Macht. Frankfurt/Main & New York: Campus Verlag.
- ⁷ Vgl. Herman, Burt (2025): Forget “Google Zero.” We need to talk about “People Zero.” In: Nieman Lab, 02.12.2025, <https://www.niemanlab.org/2025/12/forget-google-zero-we-need-to-talk-about-people-zero/>
- ⁸ Der Beauftragte der Bundesregierung für Kultur und Medien (2025): Erster virtueller Staatsminister. In: kulturstaatsminister.de, 24.09.2025, <https://kulturstaatsminister.de/erster-virtueller-staatsminister>
- ⁹ Vgl. Gerken, Tom (2023): New York Times sues Microsoft and OpenAI for ‘billions’. In: BBC.com, 27.12.2023, <https://www.bbc.com/news/technology-67826601>; Grynbaum, Michael M./Mac, Ryan (2023): The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. In: The New York Times, 27.12.2023, <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>; Brittain, Blake (2025): Judge explains order for New York Times in OpenAI copyright case. In: Reuters, 04.04.2025, <https://www.reuters.com/legal/litigation/judge-explains-order-new-york-times-openai-copyright-case-2025-04-04/>; Metz, Cade/Grynbaum, Michael M. (2025): New York Times Sues A.I. Start-Up Perplexity Over Use of Copyrighted Work. In: New York Times, 05.12.2025, <https://www.nytimes.com/2025/12/05/technology/new-york-times-perplexity-ai-law-suit.html>; ChatGPT Is Eating the World (2025): Latest map of copyright suits v. AI companies: 2 cases tentatively settle. In: ChatGPT Is Eating the World, 27.08.2025, <https://chatgptiseatingtheworld.com/2025/08/27/latest-map-of-copyright-suits-v-ai-companies-2-cases-tentatively-settle-aug-27-2025/>
- ¹⁰ Thompson, Stuart A. (2025): How Trump Is Using Fake Imagery to Attack Enemies and Rouse Supporters. In: The New York Times, 21.10.2025, <https://www.nytimes.com/interactive/2025/10/21/business/media/trump-ai-truth-social-no-kings.html>; Barnes, Renee/Riedel, Aimee/Whittaker, Lucas/Mulcahy, Rory (2024): Disinformation and deepfakes played a part in the US election. Australia should expect the same. In: The Conversation, 20.11.2024, <https://theconversation.com/disinformation-and-deepfakes-played-a-part-in-the-us-election-australia-should-expect-the-same-243373>; Myers, Steven Lee/Thompson, Stuart A. (2025): A.I. Is Starting to Wear Down Democracy. In: The New York Times, 26.06.2025, <https://www.nytimes.com/2025/06/26/technology/ai-elections-democracy.html>; Epstein, Reid J./Corasanti, Nick (2026): How Trump’s 2020 Election Claims Have Been Debunked Again and Again. In: The New York Times, 29.01.2026, <https://www.nytimes.com/2026/01/29/us/politics/trump-2020-election-claims-fact-check.html>
- ¹¹ o.A. (2025): Albanien stellt erste KI-Ministerin Europas vor. In: tagesschau.de, 23.09.2025, <https://www.tagesschau.de/ausland/europa/albanien-ki-ministerin-100.html>

- ¹² Die Bundesregierung (2025): Premiere im Kanzleramt: Staatsminister Weimer präsentiert den „Weimatar“ – Deutschlands erster Minister-Avatar stellt sich der KI-Debatte. In: bundesregierung.de, 24.09.2025, <https://www.bundesregierung.de/breg-de/aktuelles/premiere-im-kanzleramt-staatsminister-weimer-praesentiert-den-weimatar-deutschlands-erster-minister-avatar-stellt-sich-der-ki-debatte-2385478>; Der Beauftragte der Bundesregierung für Kultur und Medien (2025): Erster virtueller Staatsminister. In: kulturstaatsminister.de, 24.09.2025, <https://kulturstaatsminister.de/erster-virtueller-staatsminister>
- ¹³ AP (2026): Trump’s use of A.I. images further erodes public trust, experts say. In: pbs.org, 27.01.2026, <https://www.pbs.org/newshour/politics/trumps-use-of-ai-images-further-erodes-public-trust-experts-say>
- ¹⁴ AFP News Agency (2025): VIDEO: Albania's new AI-generated minister makes first address to parliament. In: X.com, 18.09.2025, <https://x.com/AFP/status/1968700801305477550>
- ¹⁵ Vgl. Bendel, Oliver (2021): Responsible AI. In: Gabler Wirtschaftslexikon, 06.09.2021, <https://wirtschaftslexikon.gabler.de/definition/responsible-ai-123232/version-384809>; Papagiannidis, Emmanouil/Mikalef, Patrick/Conboy, Kieran (2025): Responsible artificial intelligence governance: A review and research framework. In: The Journal of Strategic Information Systems, 34(2), 101885, <https://www.sciencedirect.com/science/article/pii/S0963868724000672>
- ¹⁶ Weizenbaum, Joseph (1966): ELIZA—A computer program for the study of natural language communication between man and machine. In: Communications of the ACM, 9(1), S. 36-45, <https://doi.org/10.1145/365153.365168>; Berry, David M. (2023): The limits of computation: Joseph Weizenbaum and the ELIZA chatbot. In: Weizenbaum Journal of the Digital Society, 3(3), <https://doi.org/10.34669/WJ.WJDS/3.3.2>; Weizenbaum Institute for the Networked Society (2026): Der ELIZA-Effekt, <https://jw.weizenbaum-institut.de/wp03>; Überland, Laf (2016): 50 Jahre Software ELIZA: Vom Psychiater inspirierte Illusion. In: Deutschlandfunk Kultur, 19.01.2016, <https://www.deutschlandfunkkultur.de/50-jahre-software-eliza-vom-psychiater-inspirierte-100.html>; Trauner, Martin (2023): KI-Pionier Joseph Weizenbaum: ELIZA und die Tücken der Ethik. In: Bayern 2 – Radio Wissen, 17.08.2025, <https://www.br.de/radio/bayern2/service/manuskripte/radiowissen/radiowissen-manuskripte-ki-pionier-joseph-weizenbaum-eliza-und-tuecken-ethik-wissenschaft-100.html>
- ¹⁷ Weichert, Stephan (2026): Die Gesellschaft des Verschwindens. In: taz, 20.02.2026, <https://taz.de/Kuenstliche-Intelligenz-Das-Ende-des-Internets-und-die-Entkernung-der-Demokratie/!6155291/>
- ¹⁸ Vgl. Berry, David M. (2023): The Limits of Computation: Joseph Weizenbaum and the ELIZA Chatbot. In: Weizenbaum Journal of the Digital Society, 3(3), <https://doi.org/10.34669/WJ.WJDS/3.3.2>; LI Academy (o. J.): The Story of ELIZA: The AI That Fooled the World, <https://liacademy.co.uk/the-story-of-eliza-the-ai-that-fooled-the-world/>; Natale, Simone (2021): The ELIZA Effect. In: The Oxford Handbook of the History of Artificial Intelligence. Oxford: Oxford University Press, <https://www.semanticscholar.org/paper/The-ELIZA-Effect-Natale/5d95d0ba2eeb6170933c067f96b84dd0204fed85>; Kalat, David (2024): The Nervous System and the ELIZA Effect. In: BRG, 03.06.2024, <https://www.thinkbrg.com/insights/publications/nervous-system-eliza-effect/>
- ¹⁹ Krotz, Friedrich (2020): Der behavioristisch verkümmerte Mensch der derzeitigen Digitalisierung und das Ende einer reflektierten Ethik? In: merz | Medien + Erziehung, 64(2), S. 16–23; Krotz, Friedrich (2024): Die Teilung geistiger Arbeit per Computer. Eine Kritik der digitalen Transformation. Baden-Baden: Nomos.
- ²⁰ Weizenbaum, Joseph (1978): Die Macht der Computer und die Ohnmacht der Vernunft. Frankfurt/Main: Suhrkamp; Weizenbaum, Joseph (1991): Against the Imperialism of Instrumental Reason. In: Kling, Rob (Hrsg.): Computerization and Controversy: Value Conflicts and Social Choices. San Diego u. a.: Academic Press, S. 728–742; Floridi, Luciano u.a. (2018): AI4People – An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707, <https://doi.org/10.1007/s11023-018-9482-5>
- ²¹ Mollick, Ethan/Mollick, Lilach (2023): Assigning AI: Seven Approaches for Students, with Prompts. Harvard Business School Working Paper No. 23-061,

- https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4475995; Microsoft Research (2023): Sparks of Artificial General Intelligence: Early experiments with GPT-4. In: arXiv:2303.12712 [cs.CL], 13.04.2023, <https://arxiv.org/abs/2303.12712>
- ²² Ein anschauliches Beispiel für diese „gezackte Grenze“ ist ein scheinbar triviales Kinderspiel wie „Schnickschnackschnuck“: In kontrollierten Tests kann KI Wahrscheinlichkeiten berechnen, Muster erkennen oder vergangene Züge analysieren. In der realen Spielsituation jedoch – mit unvorhersehbarem Verhalten, situativen Entscheidungen, Täuschung oder spontanen Regelbrüchen – versagen selbst leistungsfähige Systeme regelmäßig. Menschen wechseln Strategien, spielen bewusst irrational oder reagieren auf minimale soziale Signale. Genau diese Kontextualisierung entzieht sich maschineller Verarbeitung. Das Beispiel zeigt: KI scheitert nicht an Komplexität im technischen Sinn, sondern an Situationen, in denen Urteil, Intuition und situatives Handeln gefragt sind.
- ²³ Dell’Acqua, Fabrizio u.a. (2023): Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Harvard Business School Working Paper 24-013 (15.09.2023), <http://dx.doi.org/10.2139/ssrn.4573321>
- ²⁴ Herman, Burt (2025): Forget “Google Zero.” We need to talk about “People Zero.” In: Nieman Lab, 12.12.2025, <https://www.niemanlab.org/2025/12/forget-google-zero-we-need-to-talk-about-people-zero/>
- ²⁵ Verantwortung liegt derzeit fragmentiert bei Herstellern, Anbietern, Betreibern oder Nutzern (z. B. nach Produkthaftungs-, Urheber- oder künftig KI-Recht). Es gibt keine unmittelbare institutionelle Verantwortungskette, die Wirkung, Entscheidung und öffentliche Rechenschaft verbindet – so wie im Journalismus.
- ²⁶ Couldry, Nick/Mejias, Ulises A. (2019): The Costs of Connection. How Data Is Colonizing Human Life. Stanford: Stanford University Press; Van Dijck, José/Poell, Thomas/ De Waal, Martijn (2018): The Platform Society. Public Values in a Connective World. Oxford: Oxford University Press; Reckwitz, Andreas (2019): Das Ende der Illusionen. Politik, Ökonomie und Kultur in der Spätmoderne. Frankfurt/Main: Suhrkamp.
- ²⁷ Vgl. von Kessel-Doelle, Susanne (2023): Lisa KI-Ausgabe: Presserat rügt Burda-Verlagsgruppe. In: Meedia, 12.12.2023, <https://meedia.de/news/beitrag/14531-lisa-ki-ausgabe-presserat-ruegt-burda-verlagsgruppe.html>; Braun, Fabrice (2025): Qualitätsjournalismus, als hätte man ihn bei Temu bestellt. In: Uebermedien, 08.01.2025, <https://uebermedien.de/101728/qualitaetsjournalismus-als-haette-man-ihn-bei-temu-bestellt/>; Bartl, Marc (2025): "Bringt uns Häme als Branche wirklich weiter?": Debatte um KI-Fail beim Spiegel ebbt nicht ab. In: Kress, 04.11.2025, <https://kress.de/news/beitrag/151334-quot-bringt-uns-haeme-als-branche-wirklich-weiter-quot-debatte-um-ki-fail-beim-spiegel-ebbt-nicht-ab.html>; Niggemeier, Stefan (2026): So zerstört man Vertrauen. In: Süddeutsche Zeitung, 18.02.2026, <https://www.sueddeutsche.de/medien/zdf-heute-journal-ki-generierte-bilder-videos-ice-vertrauen-li.3389674>
- ²⁸ Weichert, Stephan (2026b): „Klasse gemacht!“ – Wie KI schleichend die Urteilskraft im Journalismus ersetzt. In: Übermedien, 24.02.2026, <https://uebermedien.de/114191/klasse-gemacht-wie-ki-schleichend-die-urteilskraft-im-journalismus-ersetzt/>
- ²⁹ Kramp, Leif/Weichert, Stephan (2024a): KI-Effekt: Wir dürfen das Vertrauen nicht verspielen. In: Medium Magazin, 05.2024, S. 54-57.
- ³⁰ Kramp, Leif/Weichert, Stephan (2024b): Resilienz in der digitalen Gesellschaft. Mediennutzung in Zeiten von Krisen, Kriegen und KI. Köln: Herbert von Halem Verlag, S. 45-46.
- ³¹ Ebenda, S. 47-49.
- ³² Ebenda, S. 50-53; vgl. Kramp, Leif/Weichert, Stephan (2024c): „Wir müssen uns von der Vorstellung lösen, dass Journalismus gleichbedeutend mit Content ist“. In: Journalist.de, 04.11.2024, <https://www.journalist.de/meinung/meinungen-detail/wir-muessen-uns-von-der-vorstellung-loesen-dass-journalismus-gleichbedeutend-mit-content-ist/>
- ³³ VOCER Institut für Digitale Resilienz (2026): <https://www.digitale-resilienz.org>
- ³⁴ Journalismus macht Schule (2026): <https://journalismus-macht-schule.org>

-
- ³⁵ Save Social (2025): www.savesocial.eu
- ³⁶ Daniel, Matthias/Weichert, Stephan (Hg.) (2022): Resilienter Journalismus. Wie wir den öffentlichen Diskurs widerstandsfähiger machen. Köln: Halem-Verlag.
- ³⁷ Vgl. Kramp, Leif/Weichert, Stephan (2025): Resilienz in der digitalen Gesellschaft. Mediennutzung in Zeiten von Krisen, Kriegen und KI. Köln: Halem-Verlag.
- ³⁸ Merkel, Wolfgang (2023): Im Zwielficht. Zerbrechlichkeit und Resilienz der Demokratie im 21. Jahrhundert. Frankfurt am Main: Campus; Merkel, Wolfgang (2024): Wie resilient ist unsere Demokratie? In: Aus Politik und Zeitgeschichte 74(27/2024), S. 18-25.
- ³⁹ Atteneder, Helena u.a. (2017): Digitale Resilienz und soziale Verantwortung. Überlegungen zur Entwicklung eines Konzepts. In: MedienJournal, 41(1), S. 48-55.
- ⁴⁰ Hepp, Andreas/Hasebrink, Uwe (2017): Kommunikative Figurationen. Ein konzeptioneller Rahmen zur Erforschung kommunikativer Konstruktionsprozesse in Zeiten tiefgreifender Mediatisierung. In: Medien & Kommunikationswissenschaft, 65, S. 330-347.
- ⁴¹ Bayerischer Rundfunk (2024): Umgang mit Künstlicher Intelligenz: Unsere KI-Richtlinien im Bayerischen Rundfunk, <https://www.br.de/extra/ai-automation-lab/ki-ethik-100.html>
- ⁴² Jones, Bronwyn (2025): Responsible AI and Public Service Journalism. BBC Research & Development, 17.10.2025, <https://www.bbc.co.uk/rd/articles/2025-10-journalism-ai-news-public-media/>
- ⁴³ The Associated Press (2024): Stylebook. 57th Edition, 2024-2026, S. 385-392.
- ⁴⁴ Wessel, Rhea (2025): What the Financial Times' AI Strategy Teaches All Organizations About Quality, Trust, and Innovation. In: Institute for Thought Leadership, 21.05.2025, <https://www.institute-forthoughtleadership.com/what-the-financial-times-ai-strategy-teaches-all-organizations-about-quality-trust-and-innovation>
- ⁴⁵ Financial Times (2025): Evolving for the AI era: How the Financial Times is creating a culture of AI fluency and innovation, <https://aboutus.ft.com/careers/ft-blog/evolving-for-the-ai-era-how-the-financial-times-is>; FT Strategies (2024): AI in the media: moving from strategy to implementation, <https://www.ftstrategies.com/en-gb/insights/ai-in-media-moving-from-strategy-to-implementation>
- ⁴⁶ Guaglione, Sara (2025): Inside The New York Times' AI newsroom strategy. In: Digiday, 22.09.2025, <https://digiday.com/media/inside-the-new-york-times-ai-newsroom-strategy/>
- ⁴⁷ Weichert, Stephan (2024a): Bitte recht menschlich. In: Süddeutsche Zeitung, 03.05.2024, <https://www.sueddeutsche.de/medien/ki-verlage-journalismus-1.6872222>.
- ⁴⁸ Roy, Nikita (2025): Tav Klitgaard: How Zetland turned a newsroom problem into a global AI business. In: Newsroom Bots, Episode 77, 16.12.2025, <https://shows.acast.com/newsroomrobots/episodes/tav-klitgaard-how-zetland-turned-a-newsroom-problem-into-a-g>
- ⁴⁹ Loewinger, Micah (2023): Gannett's AI-generated sports articles show the limits of automation in local news. In: On the Media, 08.12.2023, <https://www.wnycstudios.org/podcasts/otm/segments/story-behind-gannetts-ai-debacle-on-the-media2>
- ⁵⁰ Dupré, Maggie Harrison (2023): USA Today Updates Every AI-Generated Sports Article to Correct "Errors". In: Futurism, 06.09.2023, <https://futurism.com/usa-today-updates-ai-generated-sports>
- ⁵¹ Duffy, Clare (2023): Gannett to pause AI experiment after botched high school sports articles. In: CNN, 31.08.2023, <https://edition.cnn.com/2023/08/30/tech/gannett-ai-experiment-paused>; Wu, Daniel (2023): Gannett halts AI-written sports recaps after readers mocked the stories. In: Washington Post, 31.08.2023, <https://www.washingtonpost.com/nation/2023/08/31/gannett-ai-written-stories-high-school-sports/>
- ⁵² Vgl. Wasdahl, Alexander/Srinivasan, Ramesh (2026): Algorithms and Authors: How generative AI is transforming news production. In: First Monday, <https://firstmonday.org/ojs/index.php/fm/article/view/14442/12141>
- ⁵³ Robertson, Katie (2024): Axios Sees A.I. Coming, and Shifts Its Strategy. In: New York Times, 11.04.2024, <https://www.nytimes.com/2024/04/11/business/media/axios-ai-strategy.html>

- ⁵⁴ Owen, Laura Hazard (2025): OpenAI will fund four Axios Local newsrooms as part of a broader partnership focused on “juicing local news”. In: Nieman Lab, 15.01.2025, <https://www.niemanlab.org/2025/01/openai-will-fund-four-axios-local-newsrooms-as-part-of-a-broader-partnership-focused-on-juicing-local-news/>
- ⁵⁵ Weichert, Stephan (2024b): Bitte recht menschlich. In: Süddeutsche Zeitung, 03.05.2024, <https://www.sueddeutsche.de/medien/ki-verlage-journalismus-1.6872222>
- ⁵⁶ Blanke, Philipp (2025): Artikel nur noch mit KI: Bei Axel Springer ist die Künstliche Intelligenz nun überall. In: Der Tagesspiegel, 31.08.2025, <https://www.tagesspiegel.de/gesellschaft/medien/-ki/artikel-nur-noch-mit-ki-bei-axel-springer-ist-die-kuenstliche-intelligenz-nun-uberall-14254082.html>
- ⁵⁷ Vgl. Bartl, Marc (2025): KI als Jobkiller? Journalistenverband stellt sich gegen Axel Springer. In: Kress, 02.09.2025, <https://kress.de/news/beitrag/151053-ki-als-jobkiller-journalistenverband-stellt-sich-gegen-axel-springer.html>; Karle, Roland (2025): Springer macht Chefs zu Avataren: KI-Videos mit Poschardt und Burgard. In: Newsroom.de, 17.06.2025, <https://www.newsroom.de/news/aktuelle-meldungen/vermischtes-3/springer-macht-chefs-zu-avataren-ki-videos-mit-poschardt-und-burgard-972939/>
- ⁵⁸ Axel Springer (2025): Olaf Gersemann wird stellvertretender Chefredakteur von WELT für Künstliche Intelligenz. Pressemitteilung, 19.06.2025, <https://www.axelspringer.com/de/ax-press-release/olaf-gersemann-wird-stellvertretender-chefredakteur-von-welt-fuer-kuenstliche-intelligenz>
- ⁵⁹ Kreye, Andrian (2023): Döpfners künstliche Argumente. In: Süddeutsche Zeitung, 20.06.2023, <https://www.sueddeutsche.de/medien/springer-ki-jobabbau-1.5950435?reduced=true>
- ⁶⁰ National Public Radio (2023): Special Section: Generative Artificial Intelligence, <https://www.npr.org/about-npr/1205385162/special-section-generative-artificial-intelligence>
- ⁶¹ Ege, Konrad (2025): Öffentlicher US-Rundfunk unter Druck: Sender vor ungewisser Zukunft. In: Epd Medien, 25.08.2025, <https://medien.epd.de/article/3499>
- ⁶² Kreye, Andrian (2025): Wie die SZ im Jahr 2025 künstliche Intelligenz einsetzt, 18.07.2025, <https://www.sueddeutsche.de/kolumne/sz-ki-einsatz-journalismus-li.3218103>
- ⁶³ SRF (2026): Publizistische Leitlinien. Version 1.5, 30.01.2026, <https://publizistische-leitlinien.srf.ch/leitlinie/umgang-mit-kuenstlicher-intelligenz-ki/>
- ⁶⁴ Bauder, David (2025): Don’t like a columnist’s opinion? Los Angeles Times offers an AI-generated opposing viewpoint. In: AP.com, 16.04.2025, <https://apnews.com/article/los-angeles-times-ai-opinion-insights-ca632edd01d836ec26d27e47c0169927>
- ⁶⁵ Owen, Laura Hazard (2025): The L.A. Times adds AI-generated counterpoints to its opinion pieces and guess what, there are problems. In: Nieman Lab, 04.03.2025, <https://www.niemanlab.org/2025/03/the-l-a-times-adds-ai-generated-counterpoints-to-its-opinion-pieces-and-guess-what-there-are-problems/>
- ⁶⁶ Deck, Andrew (2025): Politico’s recent AI experiments shouldn’t be subject to newsroom editorial standards, its editors testify. In: Nieman Lab, 12.08.2025, <https://www.niemanlab.org/2025/08/politico-s-recent-ai-experiments-shouldnt-be-subject-to-newsroom-editorial-standards-its-editors-testify/>
- ⁶⁷ Krebbers, Martin (2025): Studio 47 in Duisburg: Wenn die KI eine ganze Fernsehsendung macht. In: Deutschlandfunk, 13.01.2025, <https://www.deutschlandfunk.de/studio47-in-duisburg-wenn-die-ki-eine-ganze-fernsehsendung-macht-dlf-5b3cdeb2-100.html>
- ⁶⁸ Studio 47 (2025): Richtlinien für den verantwortungsvollen Einsatz von KI im journalistischen Alltag bei Studio 47, https://www.studio47.de/services/files/sonstiges/S47_KI_Richtlinien_0225.pdf
- ⁶⁹ DJV-NRW (2024): Studio 47 produziert Newsmagazine per KI. In: Journal, 2, <https://journal-nrw.de/studio-47-produziert-newsmagazine-per-ki/>
- ⁷⁰ Orland, Kyle (2023): These AI-generated news anchors are freaking me out. In: ars technica, 15.12.2023, <https://arstechnica.com/ai/2023/12/these-ai-generated-news-anchors-are-freaking-me-out/>; Weichert, Stephan (2024c): KI im Journalismus: Roboterreporter trifft Bürgermeisterin. In: Der Tagesspiegel, 11.06.2024, <https://www.tagesspiegel.de/kultur/ki-im-journalismus-roboterreporter-trifft->

- [burgermeisterin-11778560.html](#); Weichert, Stephan (2024d): Verkaufen Redaktionen ihre Seele? In: Journalist, Serie „Mein Blick auf den Journalismus“, Juni 2024, S. 54–57, <https://www.journalist.de/werkstatt/werkstatt-detail/verkaufen-redaktionen-ihre-seele>
- ⁷¹ De-Lima-Santos, Marc-Felipe u.a. (2024): Guiding the Way: A Comprehensive Examination of AI Guidelines in Global Media. In: arXiv:2405.04706 [cs.CY], 07.05.2024, <https://doi.org/10.48550/arXiv.2405.04706>
- ⁷² Voinea, Dan Valeriu (2025): Reconceptualizing Gatekeeping in the Age of Artificial Intelligence: A Theoretical Exploration of Artificial Intelligence-Driven News Curation and Automated Journalism. In: Journalism and Media, 62(2), <https://doi.org/10.3390/journalmedia6020068>
- ⁷³ Abdulrauf, Habeeb u.a. (2025): Artificial Intelligence in Journalism: A Narrative Review of Opportunities, Challenges, Ethical Tensions, and Human-Machine Collaboration. In: American Journal of Arts and Human Science (AJAHS), 4(4), <https://doi.org/10.54536/ajahs.v4i4.5963>
- ⁷⁴ Vgl. Karlsson, Michael/Conill, Raul Ferrer/Örnebring, Henrik (2023): Recoding Journalism: Establishing Normative Dimensions for a Twenty-First Century News Media. In: Journalism Studies, 24(5), <https://doi.org/10.1080/1461670X.2022.2161929>
- ⁷⁵ Vgl. Weichert, Stephan (2025); Weichert, Stephan (2026): Die Matrix ist überall. In: Falter, 6/26, S. 22–23
- ⁷⁶ BBC (2025): Guidance: The use of Artificial Intelligence, Updated 24 June 2025, <https://www.bbc.co.uk/editorialguidelines/guidance/use-of-artificial-intelligence/#3editorialissuesint-houseofai>
- ⁷⁷ Paris Charter on AI in Journalism (2023), <https://rsf.org/sites/default/files/medias/file/2023/11/Paris%20Charter%20on%20AI%20and%20Journalism.pdf>
- ⁷⁸ World Association of News Publishers (2023): Charter for AI and Journalism: tech standards must be the responsibility of publishers. In: wan-ifra.org, 10.11.2023, <https://wan-ifra.org/2023/11/wan-ifra-steps-back-from-endorsing-new-charter-for-ai-and-journalism/>
- ⁷⁹ Netzwerk Recherche e. V. (2025): Künstliche Intelligenz (KI). Positionspapier, <https://netzwerkrecherche.org/wp-content/uploads/2025/09/KI-Positionspapier.pdf>; Deutscher Journalisten-Verband (2023): Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus, https://www.djv.de/fileadmin/user_upload/DJIV/INFORMATIONEN/medienpolitik/DJIV-Positionspapier_KI_2023-04.pdf; Deutscher Journalisten-Verband (2024): Stellungnahme des Deutschen Journalisten-Verbandes e.V. zu KI und den möglichen Auswirkungen auf den Journalismus anlässlich der Einladung des Wirtschaftsforums der SPD, https://www.djv.de/fileadmin/user_upload/DJIV/INFORMATIONEN/medienpolitik/05012024_DJIV-StN_zu_KI_Wirtschaftsforum_SPD_final.pdf
- ⁸⁰ ARD/Deutsche Welle/Deutschlandradio/ZDF (2026): KI-Kodex, <https://www.ard.de/die-ard/presse-und-kontakt/ard-presse-meldungen/2026/KI-Kodex-100.pdf>
- ⁸¹ Auf diese Problematik verweist auch eine internationale Studie der Europäischen Rundfunkunion (EBU) in Kooperation mit der BBC (vgl. EBU/BBC: News Integrity in AI Assistants. An international PSM study, Oktober 2025, https://www.ebu.ch/Report/MIS-BBC/NI_AI_2025.pdf). Sie zeigt, dass KI-Systeme journalistische Inhalte verzerren – unabhängig von Sprache, Land oder Mediengattung. Nicht spektakulär, nicht als offensichtliche Fehler, sondern als schleichende Normalität: durch Glättung, Verkürzung, falsche Kontextualisierung oder scheinbar plausible Umformulierungen.
- ⁸² Bayerischer Rundfunk (2020/2024): Umgang mit Künstlicher Intelligenz: Unsere KI-Richtlinien im Bayerischen Rundfunk, <https://www.br.de/extra/ai-automation-lab/ki-ethik-100.html>
- ⁸³ Holzki, Larissa (2025): Panne beim „Spiegel“ – Passen KI und Journalismus zusammen? In: Handelsblatt, 25.10.2025, <https://www.handelsblatt.com/technik/ki/ki-briefing-panne-beim-spiegel-passen-ki-und-journalismus-zusammen/100167284.html>
- ⁸⁴ Der Spiegel (2025): In eigener Sache: Wie der Spiegel künstliche Intelligenz einsetzt, <https://www.spiegel.de/backstage/in-eigener-sache-wie-der-spiegel-kuenstliche-intelligenz-einsetzt-a-146d6a53-fb8b-4d0f-8e88-0ae02f64c06a>

- ⁸⁵ Kreye, Andrian (2025): Wie die SZ im Jahr 2025 künstliche Intelligenz einsetzt, In: Süddeutsche Zeitung, 18.07.2025, <https://www.sueddeutsche.de/kolumne/sz-ki-einsatz-journalismus-li.3218103>
- ⁸⁶ Kilg, Marie u.a. (2022): Kolumne einer künstlichen Intelligenz. Wie schreibt eine Robo-Autor*in? In: taz, 26.11.2022, <https://taz.de/Kolumne-einer-kuenstlichen-Intelligenz!/5898282/>; Salzer, Robert u.a. (2024): Das Ende einer KI-Kolumne. Hinterher sind wir alle klüger. In: taz, 15.04.2024, <https://taz.de/Das-Ende-einer-KI-Kolumne!/6001264/>
- ⁸⁷ taz – die tageszeitung (2025): Wie die taz Künstliche Intelligenz nutzt: Unsere KI-Leitlinien (Update: Juli 2025), <https://taz.de/Wie-die-taz-Kuenstliche-Intelligenz-nutzt!/vn5967635/>
- ⁸⁸ Vgl. Deutsche Presse-Agentur (2023): Offen, verantwortungsvoll und transparent – Die Guidelines der dpa für Künstliche Intelligenz, <https://innovation.dpa.com/2023/04/03/kuenstliche-intelligenz-fuenf-guidelines-der-dpa/>
- ⁸⁹ Deutsche Presse-Agentur (2025): Wegweiser KI. KI-Handreichung für Redaktionen, <https://epa-per.dpa.com/issue.act?issueld=247734&issueMutation=whitepaper&issueDate=20250408®ion=DPA>
- ⁹⁰ Vgl. Kramp, Leif/Backhaus, Katia (2025): „Wir zehren noch vom Image vergangener Tage“. In: Journalist, 11, November 2025, S. 24-28.
- ⁹¹ Deutsche Forschungsgemeinschaft (DFG) (2023): Stellungnahme des Präsidiums zu Künstlicher Intelligenz, Daten und Forschung. Bonn; acatech – Deutsche Akademie der Technikwissenschaften (2024): Künstliche Intelligenz verantwortungsvoll nutzen. Impulspapier: München.
- ⁹² Vgl. Bundesärztekammer (2025): Thesenpapier KI in der Gesundheitsversorgung, https://www.bundesaeztekammer.de/fileadmin/user_upload/BAEK/Themen/Digitalisierung/Thesenpapier_KI_in_der_Gesundheitsversorgung_03.2025.pdf
- ⁹³ Vgl. Bundesministerium des Innern (BMI) (2025): Leitlinien für den Einsatz von Künstlicher Intelligenz in der Bundesverwaltung, https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/moderne-verwaltung/ki/BMI25020-leitlinien-ki-bundesverwaltung.pdf?__blob=publicationFile&v=4; Bundesministerium für Arbeit und Soziales (2025): Selbstverpflichtende Leitlinien für den KI-Einsatz in der behördlichen Praxis der Arbeits- und Sozialverwaltung, https://www.bmas.de/Shared-Docs/Downloads/DE/Publikationen/a862-01-leitlinien-ki-einsatz-behoerdliche-praxis-arbeits-sozialverwaltung.pdf?__blob=publicationFile&v=2
- ⁹⁴ UNESCO (2023): Guidance on Generative AI in Education and Research, <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- ⁹⁵ European Broadcasting Union (2025): News Integrity in AI Assistants. October 2025, https://www.ebu.ch/files/live/sites/ebu/files/Publications/MIS/open/EBU-MIS-BBC_News_Integrity_in_AI_Assistants_Report_2025.pdf
- ⁹⁶ European Commission (2024): Artificial Intelligence Act – Consolidated Version, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- ⁹⁷ Vgl. Endnote 9.
- ⁹⁸ Vgl. Cigref/Asterès (2025): Technological Dependence on American Software and Cloud Services: An Assessment of the Economic Consequences in Europe. Paris: Cigref, April 2025, S. 18, <https://www.cigref.fr/wp/wp-content/uploads/2025/05/TECHNOLOGICAL-DEPENDENCE-ON-AMERICAN-SOFTWARE-AND-CLOUD-SERVICES-AN-ASSESSMENT-OF-THE-ECONOMIC-CONSEQUENCES.pdf>
- ⁹⁹ Züger, Theresa/Asghari, Hadi (2023): AI for the public. How public interest theory shifts the discourse on AI. In: AI & Society, 38, <https://doi.org/10.1007/s00146-022-01480-5>
- ¹⁰⁰ Vgl. Bozeman, Barry (2007): Public values and public interest counterbalancing economic individualism. Washington, D.C.: Georgetown University Press, S. 12.
- ¹⁰¹ Vgl. Schwäbe, Carsten/Hummeler, Andreas/Blind, Knut (2024): Exploring an Innovation Policy for Public AI – Rationales, Examples and Learning. Policy brief of the project “Public AI”. Karlsruhe: Fraunhofer ISI, <https://doi.org/10.24406/publica-3989>

-
- ¹⁰² Vgl. Züger, Theresa/Asghari, Hadi (2023): AI for the public. How public interest theory shifts the discourse on AI. In: AI & Society, 38, <https://doi.org/10.1007/s00146-022-01480-5>, S. 822f.
- ¹⁰³ LAION (2023): A Call to protect Open-Source AI in Europe, <https://laion.ai/notes/letter-to-the-eu-parliament/>
- ¹⁰⁴ Vgl. Tiwari, Udbhav (2024): The Openness Imperative: Charting a Path for Public AI. In: AI Now Institute (Hrsg.): Redirecting Europe's AI Industrial Policy. From Competitiveness to Public Interest. New York: AI Now Institute, S. 59-63, <https://ainowinstitute.org/wp-content/uploads/2024/10/IX.-The-Openness-Imperative-Charting-a-Path-for-Public-AI.pdf>
- ¹⁰⁵ European Commission (2024): Artificial Intelligence Act – Consolidated Version, Annex XI, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- ¹⁰⁶ Current AI (2025): Building Public Interest AI - Current AI's Next Chapter. <https://www.current-ai.org/blogs/building-public-interest-ai---current-ais-next-chapter>
- ¹⁰⁷ Tarkowski, Alek/Sieker, Felix (2026): Public AI Policy Brief. Gütersloh: Bertelsmann-Stiftung, <https://doi.org/10.11586/2026001>
- ¹⁰⁸ Ebenda.
- ¹⁰⁹ AIAAIC Repository (2025): LAION-5B dataset, <https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/laion-5b-dataset>
- ¹¹⁰ Vgl. Knibbs, Kate (2024): Publishers Target Common Crawl In Fight Over AI Training Data. In: Wired, 12.06.2024, <https://www.wired.com/story/the-fight-against-ai-comes-to-a-foundational-data-set/>
- ¹¹¹ Tarkowski, Alek (2024): AI Act fails to set meaningful dataset transparency standards for open source AI. In: Open Future, 07.03.2024. <https://openfuture.eu/blog/ai-act-fails-to-set-meaningful-dataset-transparency-standards-for-open-source-ai/>; Gahntz, Maximilian (2023): The EU's AI Act and Foundation Models: The Final Stretch. In: Mozilla Foundation, 10.08.2023, <https://www.mozillafoundation.org/en/blog/the-eus-ai-act-and-foundation-models-the-final-stretch/>
- ¹¹² Keller, Paul (2023): A Frankenstein-Like Approach: Open Source in the AI Act. In: Open Future, 14.12.2023, <https://openfuture.eu/blog/a-frankenstein-like-approach-open-source-in-the-ai-act/>
- ¹¹³ Böcker, Jens (2023): Implementierungsstrategien für die Gaia-X Föderationsdienste: Projektfortschritte und neue Anforderungen eine Delta-Studie mit Ausblick. GXSF Studie, <https://web.archive.org/web/20240216154612/https://www.gxfs.eu/download/10152>
- ¹¹⁴ Vgl. Fermigier, Stefane (2025): Gaia-X – Chronicle of a Failure Foretold. In: euro-stack.com, 04.02.2025, <https://euro-stack.com/blog/2025/2/gaia-x-failure>
- ¹¹⁵ Forum on Information and Democracy (2024): A Voluntary Certification Mechanism for Public Interest AI. Research Paper, <https://informationdemocracy.org/wp-content/uploads/2024/09/FID-Public-Interest-AI-Sept-2024.pdf>
- ¹¹⁶ Wikimedia/Gewerkschaft Erziehung und Wissenschaft (GEW) (2024): 10 Empfehlungen für die Bildungs- und Digitalpolitik. In: GEW Aktuelles, 15.05.2024, <https://www.gew.de/aktuelles/detailseite/10-empfehlungen-fuer-eine-offene-ki-in-der-paedagogik>
- ¹¹⁷ Equitech Futures (2025): Reclaiming AI's Promise. A Roadmap for Public Interest AI. Report, https://cdn.prod.website-files.com/67a498f10e7c0758d8cd4650/685bfc52036aefa872c9896e_2fea60182d428c528549670a740c3b16_Equitech%20Futures%20-%20Public%20Interest%20AI.pdf
- ¹¹⁸ Züger, Theresa/Asghari, Hadi/Lübbert, Birte (2025): Public Interest AI Projects 2024. In: Huggingface, 30.01.2025, <https://huggingface.co/datasets/hiig-ai-lab/publicinterest-ai-projects>
- ¹¹⁹ European Commission (2025): The General-Purpose AI Code of Practice, <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>; Code of Practice AI (2025): EU AI Act: General-Purpose AI Code of Practice – Final Version. <https://code-of-practice.ai>
- ¹²⁰ Vgl. Belitz, Oliver/Hemmbt, Simon (2026): Taking the EU AI Act to Practice: Understanding the Draft Transparency Code of Practice. In: Bird & Bird, 25.01.2026, <https://www.twobirds.com/en/insights/2026/taking-the-eu-ai-act-to-practice-understanding-the-draft-transparency-code-of-practice>

- ¹²¹ Bundesministerium für Digitales und Verkehr/Bundesnetzagentur (2024): Zuständigkeit für die Durchsetzung der KI-Verordnung. https://www.bundesnetzagentur.de/DE/Fachthemen/Digitales/KI/start_ki.html; Bundesministerium für Bildung und Forschung/BMDS (2025): Gesetz zur Durchführung der KI-Verordnung, 12.09.2025, <https://bmbs.bund.de/service/gesetzgebungsverfahren/gesetz-zur-durchfuehrung-der-ki-verordnung>
- ¹²² Bundesverband Digitalpublisher und Zeitungsverleger (BDZV)/Medienverband der Freien Presse (MVFP) (2023): Einigung zum KI-Gesetz der EU: Eine Enttäuschung für Medien- und Kulturschaffende, 11.12.2023, <https://www.bdzv.de/service/presse/pressemitteilungen/2023/einigung-zum-ki-gesetz-der-eu-eine-enttaeuschung-fuer-medien-und-kulturschaffende>; BDZV/Retresco (2025): KI-Reifegrad-Report 2025, <https://www.bdzv.de/service/presse/branchennachrichten/2025/ki-in-den-redaktionen-2025-viel-einsatz-noch-wenig-wirkung>
- ¹²³ Vgl. epd medien (2025): BDZV fordert weitergehende Regelungen zum Umgang mit KI, <https://medien.epd.de/article/3458>
- ¹²⁴ Ver.di (2024): Tarifverhandlung Zeitungsredaktionen – ohne Ergebnisse zu Tarifierhöhungen und KI-Regeln, <https://www.verdi.de/presse/pressemitteilungen/tarifverhandlung-zeitungsredaktionen-ohne-ergebnisse-tarifierhoehungen-und-ki-regeln>
- ¹²⁵ UNESCO (2025): World Trends in Freedom of Expression and Media Development 2022–2025. Journalism: Shaping a World at Peace, <https://unesdoc.unesco.org/ark:/48223/pf0000396638>
- ¹²⁶ Eine gemeinsame Erklärung der Beauftragten für Meinungsfreiheit der Vereinten Nationen, der OSZE, der OAS und der Afrikanischen Menschenrechtskommission vom Oktober 2025 macht zugleich deutlich: KI-Systeme können Pressefreiheit und freie Informationsordnung stärken – oder untergraben. Diese Verantwortung trägt nicht eine Institution allein. Staaten, Aufsichtsbehörden, Plattform- und KI-Unternehmen, Medienorganisationen, Zivilgesellschaft und Wissenschaft seien gemeinsam gefragt: OAS/United Nations/OSCE/ ACHPR (2025): Joint Declaration on AI, Freedom of Expression and Media Freedom, 24.10.2025, <https://rfom.osce.org/representative-on-freedom-of-media/600217>
- ¹²⁷ Weichert, Stephan (2026): „Klasse gemacht!“ – Wie KI schleichend die Urteilskraft im Journalismus ersetzt. In: Übermedien, 24.02.2026, <https://uebermedien.de/114191/klasse-gemacht-wie-ki-schleichend-die-urteilskraft-im-journalismus-ersetzt/>
- ¹²⁸ O'Connor, Hailey (2024): How Journalistic Red-Teaming Can Help AI Companies Fix Misinformation Blindspots. In: NewsGuard, Oktober 2024, <https://www.newsguardtech.com/insights/how-journalistic-red-teaming-can-help-ai-companies-fix-misinformation-blindspots/>
- ¹²⁹ The Guardian: (2023): The Guardian's approach to generative AI. In: The Guardian, 16.06.2023, <https://www.theguardian.com/help/insideguardian/2023/jun/16/the-guardians-approach-to-generative-ai>; Sénécat, Adrien: (2024): 80 years of Le Monde: Your questions answered about how our newsroom works. In: Le Monde, 19.12.2024, https://www.lemonde.fr/en/les-decodeurs/article/2024/12/19/80-years-of-le-monde-your-questions-about-how-our-newsroom-works-answered_6736236_8.html; The Washington Post (2024): Ask the Post AI. In: The Washington Post, <https://www.washingtonpost.com/ask-the-post-ai/>; The Washington Post (2024): About Ask the Post AI. In: The Washington Post, 7.11.2024 (Update: 24.11.2025), <https://www.washingtonpost.com/technology/2024/11/07/faq-ask-the-post-ai/>; The Washington Post (2024): The Washington Post launches "Ask The Post AI," a new search experience. In: The Washington Post, 7.11.2024, <https://www.washingtonpost.com/pr/2024/11/07/washington-post-launches-ask-post-ai-new-search-experience/>; ProPublica (o.J.): Do You Have a Tip for ProPublica? Here's How to Contact Us Securely. In: ProPublica, <https://www.propublica.org/tips/>
- ¹³⁰ BBC (2023): BBC News puts transparency at its heart with BBC Verify, 17.05.2023, <https://www.bbc.co.uk/mediacentre/bbc-news-transparency-bbc-verify>
- ¹³¹ Schultz, Ray (2023): 'The Guardian' Updates Its Editorial Code. In: MediaPost, 27.07.2023, <https://www.mediapost.com/publications/article/387650/the-guardian-updates-its-editorial-code.html?edition=131107>

- ¹³² SRG Deutschschweiz (2024): SRF führt KI-Richtlinien ein, <https://www.srgd.ch/de/aktuelles/news/srf-passt-publizistische-leitlinien-an-und-fuhrt-ki-richtlinien-ein/>
- ¹³³ Harrison Dupré, Maggie (2023): Sports Illustrated Published Articles by Fake, AI-Generated Writers. In: Futurism, 27.11.2023, <https://futurism.com/sports-illustrated-ai-generated-writers>; Napolitano, Elizabeth (2023): Sports Illustrated owner denies using AI and fake writers to produce articles. In: CBS News, 28.11.2023, <https://www.cbsnews.com/news/sports-illustrated-denies-using-ai-fake-writers-to-produce-stories/>; Halpert, Madeline (2023): Sports Illustrated publisher fires CEO Ross Levinsohn after AI scandal. In: BBC.com, 12.12.2023, <https://www.bbc.com/news/world-us-canada-67619015>
- ¹³⁴ Bonifacic, Igor (2023): CNET had to correct most of its AI-written articles. In: Engadget, 25.01.2023, <https://sg.news.yahoo.com/cnet-corrected-41-of-its-77-ai-written-articles-201519489.html>
- ¹³⁵ Sato, Mia/Roth, Emma (2023): CNET found errors in more than half of its AI-written stories. In: The Verge, 25.01.2023, <https://www.theverge.com/2023/1/25/23571082/cnet-ai-written-stories-errors-corrections-red-ventures>
- ¹³⁶ The New York Times Company (2024): Principles for Using Generative A.I. in The Times's Newsroom, 09.05.2024, <https://www.nytc.com/press/principles-for-using-generative-a%E2%80%A4%E2%80%A4-in-the-timess-newsroom/>
- ¹³⁷ BBC (2023): BBC News puts transparency at its heart with BBC Verify, 17.05.2023, <https://www.bbc.co.uk/mediacentre/bbc-news-transparency-bbc-verify>
- ¹³⁸ Mahadevan, Alex (2023): Associated Press cements the AI era with newsroom guidance. In: Poynter.org, 16.08.2023, <https://www.poynter.org/ethics-trust/2023/new-ap-stylebook-guidelines-artificial-intelligence-chatgpt/>
- ¹³⁹ Vgl. Neuts, Dana E. (2023): AP Prohibits Use of Generative AI for News Stories or Images. In: Subscription Insider, 21.08.2023, <https://www.subscriptioninsider.com/article-type/news/ap-prohibits-use-of-generative-ai-for-news-stories-or-images>
- ¹⁴⁰ Leffer, Lauren (2023): CNET Is Reviewing the Accuracy of All Its AI-Written Articles After Multiple Major Corrections. In: Gizmodo, 17.01.2023, <https://gizmodo.com/cnet-ai-chatgpt-news-robot-1849996151>
- ¹⁴¹ Carnevale, Robert (2023): Gizmodo Editor Slams 'Shameful' AI-Written Article: 'It's F-king Dogs-t'. In: The Wrap, 05.07.2023, <https://www.thewrap.com/ai-gizmodo-io9-star-wars-article-retraction/>] AIAAIC (2023): Gizmodo AI generates error-strewn Star Wars article. In: AIAAIC, July 2023, <https://www.ai-aaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents/gizmodo-ai-generates-error-strewn-star-wars-article>
- ¹⁴² Bernoff, Josh (2023): AP issues guidance on AI content. Basically: "Don't publish bullshit." In: Bernoff.com, 18.08.2023, <https://bernoff.com/blog/ap-issues-guidance-on-ai-content-basically-dont-publish-bullshit>
- ¹⁴³ SRF (2026): Publizistische Leitlinien, Version 1.5, 30.01.2026, <https://publizistische-leitlinien.srf.ch>
- ¹⁴⁴ Good Tape (2026): Good Tape for journalism, <https://goodtape.io/good-tape-for-newsrooms/>
- ¹⁴⁵ Ippen Digital (2024): Newsletter: Ippen Digital führt agentische KI ein + Beat Squares will Lokalredaktionen helfen, Oktober 2024, <https://lokaljournalismus.digital/ippen-digital-fuehrt-agentische-ki-ein-beat-squares-will-lokalredaktionen-helfen-livestreaming-revolutioniert-lokaljournalismus-modernisierung-von-lokaljournalismus-durch-geodaten/>
- ¹⁴⁶ Sommer, Rupert (2025): Markus Knall und das 1.000-Artikel-Experiment bei Ippen. In: Newsroom.de, 26.06.2025, <https://www.newsroom.de/news/aktuelle-meldungen/vermishtes-3/markus-knall-und-das-1000-artikel-experiment-bei-ippen-lokaljournalismus-per-ki-skalisieren-973212/>
- ¹⁴⁷ Reuters (2023): How AI helps power trusted news at Reuters, 12.07.2023, <https://www.reuters.com/media-center/how-ai-helps-power-trusted-news-at-reuters/>
- ¹⁴⁸ Leffer, Lauren (2023): CNET Is Reviewing the Accuracy of All Its AI-Written Articles After Multiple Major Corrections. In: Gizmodo, 17.01.2023, <https://gizmodo.com/cnet-ai-chatgpt-news-robot-1849996151>

-
- ¹⁴⁹ Vgl. VandeHei, Jim (2025): Axios, OpenAI go big in local. In: Axios, 15.01.2025, <https://www.axios.com/2025/01/15/axios-openai-go-big-in-local>;
- ¹⁵⁰ Vgl. Axios (2025): Axios, OpenAI partner to expand local news coverage, Press Release, 15.01.2025, <https://www.axios.com/press-past-releases/axios-openai-partner-to-expand-local-news-coverage>
- ¹⁵¹ Christian, Jon (2023): Magazine Publishes Serious Errors in First AI-Generated Health Article. In: Futurism, 18.02.2023, <https://futurism.com/neoscope/magazine-mens-journal-errors-ai-health-article>
- ¹⁵² Vgl. Wessel, Rhea (2025): What the Financial Times' AI Strategy Teaches All Organizations About Quality, Trust, and Innovation. In: Institute for Thought Leadership, 21.05.2025, <https://www.institute-forthoughtleadership.com/what-the-financial-times-ai-strategy-teaches-all-organizations-about-quality-trust-and-innovation>
- ¹⁵³ Vgl. Spencer, Clare (2025): How a Danish News Service Made a Profit with its Transcription Tool. In: Medium, 02.06.2025, <https://generative-ai-newsroom.com/how-a-danish-news-service-made-a-profit-with-its-transcription-tool-285bc05b7cf9>
- ¹⁵⁴ Maher, Bron (2024): Reach using AI to speed up 'ripping' and use same article on multiple sites. In: PressGazette, 22.02.2024, <https://pressgazette.co.uk/publishers/nationals/reach-ai-guten/>
- ¹⁵⁵ National Union of Journalists (2025): NUJ deeply concerned by sweeping Reach redundancies, 10.07.2025, <https://www.nuj.org.uk/resource/nuj-deeply-concerned-by-sweeping-reach-redundancies.html>
- ¹⁵⁶ Washington Post (2024): The Washington Post Launches 'Ask The Post AI,' a New Search Experience, Press Release, 07.11.2024, <https://www.washingtonpost.com/pr/2024/11/07/washington-post-launches-ask-post-ai-new-search-experience/>
- ¹⁵⁷ Vgl. Weiss, Jessica (2018): How ProPublica invites readers to get involved. In: International Journalists' Network, 30.10.2018, <https://ijnet.org/en/story/how-propublica-invites-readers-get-involved>; Tobin, Ariana/Gallardo, Adriana/Raghavendran, Beena (2019): What Engagement Reporting Does — and Doesn't — Mean at ProPublica. In: ProPublica, 15.01.2019, <https://www.propublica.org/article/what-does-engagement-reporting-mean-propublica>
- ¹⁵⁸ ProPublica (2025): ProPublica Joins Lenfest Institute AI Collaborative and Fellowship Program, 28.03.2025, <https://www.propublica.org/atpropublica/propublica-joins-lenfest-institute-ai-collaborative-and-fellowship-program>
- ¹⁵⁹ Thompson, Sarah (2025): Inside ProPublica's responsible AI strategy for investigations. In: AISetup, 13.03.2025, <https://aisetup.net/news/inside-propublicas-responsible-ai-strategy-for-investigations/>
- ¹⁶⁰ Mullin Benjamin/Robertson, Katie/Wemple, Erik (2026): Washington Post Lays Off More Than 300 Journalists. In: The New York Times, 4.02.2026, <https://www.nytimes.com/2026/02/04/business/media/washington-post-layoffs.html>

Die Autoren:

Dr. Leif Kramp ist Medien- und Journalismusforscher sowie Forschungs koordinator des Zentrums für Medien-, Kommunikations- und Informationsforschung (ZeMKI), Universität Bremen. Kramp ist Gründungsvorstand des Vereins für Medien- und Journalismuskritik e.V., der die VOCER-Bildungsprogramme trägt. In seiner Arbeit verbindet er wissenschaftliche Analyse mit praxisnaher Kooperation in Projekten mit Medienhäusern, Stiftungen und öffentlichen Institutionen. Er begleitet Transformationsprozesse in Redaktionen und Organisationen, analysiert strukturelle Veränderungen im Mediensystem und berät bei strategischen Weichenstellungen im digitalen Wandel. Kramp ist Autor und Herausgeber zahlreicher Fachbücher und Studien zum Wandel von Medien, Journalismus und Gesellschaft.



E-Mail: kramp@uni-bremen.de

Dr. Stephan Weichert ist Medien- und Kommunikationswissenschaftler, Resilienz-Coach für Führungskräfte und Strategieberater. Seit 2020 leitet er das unabhängige VOCER-Institut für Digitale Resilienz, einen gemeinnützigen Think & Do Tank, der Weiterbildung, Beratung und Forschung rund um den souveränen Umgang mit digitalen Medien sammendenkt. Weichert arbeitet seit über 25 Jahren als Lehrbeauftragter und Professor für digitalen Journalismus (u.a. Universität der Künste, City University of New York, Hamburg Media School, Macromedia University, TU Dortmund). Der Medienexperte berät u.a. Ministerien und Organisationen zu Kernfragen digitaler Nachrichten-kompetenz. Seine Arbeit wurde mehrfach ausgezeichnet, u.a. mit dem „Medienethik Award“. Aktuell lehrt und forscht er zum Thema „KI-Resilienz in der digitalen Welt“.



E-Mail: weichert@vocer.org

Impressum

Kramp, Leif/Weichert, Stephan (2026):

Whitepaper: KI-Resilienz im Journalismus. Framework und Handlungsempfehlungen für mehr Vertrauen, Verantwortung und digitale Souveränität.

Hamburg: VOCER-Institut für Digitale Resilienz.

Erscheinung: 15. März 2026

© 2026 VOCER-Institut für Digitale Resilienz

c/o Verein für Medien- und Journalismuskritik e.V.

Postfach 20 14 54 | 20204 Hamburg

info@digitale-resilienz.org | www.digitale-resilienz.org