

Aufsatz

Kim Björn Becker

Neues Spiel, neue Regeln

Eine Untersuchung von redaktionellen Richtlinien für den Umgang mit Künstlicher Intelligenz im Newsroom

Abstract: Mit der Einführung des Sprachmodells ChatGPT hat die Nutzung von Künstlicher Intelligenz (KI) einen Hype erfahren – auch im Journalismus, dessen professioneller Umgang mit Sprache breite Anwendungsfelder eröffnet. Mit den neuen Möglichkeiten entstehen für Medienorganisationen aber auch Fragen zum Umgang mit KI. Erste Redaktionen haben auf die Herausforderung mit der Veröffentlichung eigener KI-Richtlinien reagiert: Mit ihnen wollen sie klarmachen, nach welchen Prinzipien sie die Algorithmen nutzen. Die Arbeit stellt durch eine vergleichende Untersuchung der Dokumente von sieben internationalen Medien ein grundsätzliches Verständnis dafür her, in welchen Bereichen Redaktionen Chancen sehen und welche Fallstricke sie adressieren. Insgesamt wurden jeweils zwei Organisationen aus Deutschland und den Vereinigten Staaten sowie je eine aus den Niederlanden, Großbritannien und Kanada in die Untersuchung eingeschlossen. Die Auswertung zeigt, dass Nachrichtenagenturen sich eher konziser formulierte Regeln geben, während öffentlich-rechtliche Rundfunkanstalten einen umfassenderen Regulierungsanspruch verfolgen. Dabei setzen die Redaktionen unterschiedliche Schwerpunkte. Während fast alle Medien sich in ihren Leitlinien mit der Kontrolle von KI durch den Menschen sowie Fragen der Transparenz befassen, stehen Anforderungen an vertrauenswürdige Algorithmen weniger im Mittelpunkt. Die Untersuchung zeigt, dass Medien sich zwar bereits mit wesentlichen Fragen der neuen Technik beschäftigen, es in Newsrooms aber noch blinde Flecken beim Umgang mit KI gibt.

Keywords: Künstliche Intelligenz, Richtlinien, Journalistische Ethik, Autorenzeile, Newsroom

1. Einleitung: Warum Redaktionen nicht einfach nichts tun

Redaktionen nutzen Künstliche Intelligenz (KI) bereits an praktisch allen Punkten des journalistischen Prozesses (vgl. DIAKOPOULOS 2019: 76f.; BUXMANN/SCHMIDT 2022). Mit den neuen technischen Möglichkeiten bei Recherche, Produktion und Distribution entstehen für Medienorganisationen neue, drängende Fragen. Wie sollen sie mit Sprachmodellen wie ChatGPT und Bard verantwortungsvoll umgehen? Von welchen Zielen sollen sie sich beim Einsatz von KI im Newsroom leiten lassen? Und welche Gefahren gilt es dabei möglichst zu umschiffen? Weil sich inhaltliche Verzerrungen und faktische Fehler unmittelbar auf die wahrgenommene Glaubwürdigkeit der Berichterstattung auswirken dürften, brauchen Chefredaktionen und Herausgeber schnell Klarheit.

Die Standesethik bietet zwar eine gute Grundlage, um erste Antworten auf diese Fragen zu finden. Doch die journalistischen Grundsätze, die von Fachgesellschaften wie dem Deutschen Presserat und der amerikanischen Society of Professional Journalists niedergeschrieben wurden, bleiben in der Regel allgemein und für die anspruchsvollen Fragestellungen zu abstrakt. Es gibt Ausnahmen. In Brüssel hat der Raad voor de Journalistiek eine Richtlinie zur Ergänzung des belgischen Pressekodexes veröffentlicht, in der es um die Transparenz der journalistischen Arbeitsweise geht. Demnach müssen die Nutzerinnen und Nutzer wissen, wenn ein Algorithmus an einer Geschichte mitgeschrieben hat (vgl. RAAD VOOR DE JOURNALISTIEK o.J.). In Spanien hat der katalanische Presserat Redaktionen im Dezember 2021 eine Handreichung in acht Regeln gegeben, mit denen sie mögliche Fallstricke bei Datennutzung, Transparenz und algorithmischer Verzerrung umgehen können (vgl. VENTURA POCINO 2021). Und in Deutschland hat der Deutsche Journalisten-Verband (DJV) in einem Ende April veröffentlichten Positionspapier klargestellt, dass sich Redaktionen nicht aus der Verantwortung für Inhalte, an deren Zustandekommen KI beteiligt war, »stehlen« können (DEUTSCHER JOURNALISTEN-VERBAND 2023: 1). KI-Anwendungen agierten demnach »fernab von Ethik und einem Wertesystem« und seien daher »nicht in der Lage, die Wächterfunktion, die Journalistinnen und Journalisten seit jeher zukommt, zu übernehmen« (ebd.). In neun Punkten definiert der Verband sodann Leitplanken zur redaktionellen Nutzung von KI, die einen verantwortungsvollen und transparenten Umgang mit der neuen Technik sicherstellen sollen (vgl. DEUTSCHER JOURNALISTEN-VERBAND 2023: 2f.).

Doch wenn es um die konkrete Umsetzung im Newsroom geht, sind Redaktionen in Sachen KI meist auf sich gestellt – auch weil die Implementation von Algorithmen in Medienorganisationen als eine anspruchsvolle kommunikative Aufgabe gesehen werden kann (vgl. SKRUBBELTRANG MAHNKE/KARLIN 2023), für die es kaum ein Lehrbuch gibt. Erste Organisationen haben bereits darauf reagiert, indem sie sich selbst Richtlinien für den Umgang mit KI im Newsroom

und darüber hinaus gegeben haben – doch die Antworten der verantwortlichen Redakteurinnen und Redakteure fallen ganz unterschiedlich aus. In dieser Arbeit werden sieben Richtlinien internationaler Medien formal und inhaltlich untersucht. Es geht darum, ein grundlegendes Verständnis der frühen Ansätze der medialen Selbstregulierung zu gewinnen. Wie detailliert und verbindlich sind die Regeln, die sich Medienhäuser selbst geben? Auf welche Ziele und journalistischen Werte beziehen sie ihr Engagement in Sachen KI? Und muss der Mensch jede Meldung, an der ein großes Sprachmodell mitgewirkt hat, kontrollieren und freigeben? Die vorliegende Arbeit will durch die vergleichende Untersuchung von sieben regulativen Dokumenten ein grundsätzliches Verständnis dafür herstellen, welche redaktionellen Anwendungsbereiche internationale Medien für algorithmische Anwendungen grundsätzlich für geeignet halten, welche Anforderungen sie an »Responsible AI«¹⁾ stellen – also jene KI, die nach bestimmten Vorgaben entwickelt und trainiert worden ist – und wie es um die Transparenz bei der Nutzung bestellt ist.

2. Empfehlungen, Gesetzbücher und Leitfragen: Formale Aspekte von redaktionellen KI-Richtlinien

Keine redaktionelle Richtlinie zum Umgang mit KI gleicht der anderen – nicht im Inhalt, nicht in der Form. Mal kommen die Leitplanken selbst wie ein kraftvoller journalistischer Text daher, mal formalistisch und nüchtern wie die Verordnung eines Ministeriums. Manche Dokumente verbinden jede Regel mit einer ethischen Selbstvergewisserung, andere wirken selbst wie ein Algorithmus – eine Arbeitsanweisung also, die interdisziplinäre Projektteams anleiten soll. Und während das eine Dokument unumstößlich wirkende Gesetze aufstellt, formuliert ein anderes allenfalls vorsichtige Leitfragen.

Vor der inhaltlichen Auswertung gilt es, zunächst einen Blick darauf zu werfen, wie die relevanten Dokumente überhaupt beschaffen sind. Wer hat sie publiziert und wann? Welchen Anspruch verfolgen die Autorinnen und Autoren damit? Und wie detailliert sind die Vorgaben zum Umgang mit KI ausformuliert?

1 Dignum (2019: 93f.) weist darauf hin, dass mit dem Konzept meist unterschiedliche Merkmale und Anforderungen verbunden werden. Um Missverständnisse zu vermeiden, wird der Begriff hier im englischen Original verwendet und nicht in deutscher Übersetzung; AI steht dabei für Artificial Intelligence. »Verantwortungsvolle KI« wäre zwar eine korrekte Übersetzung, aber insofern unscharf, als dass eine KI nicht verantwortungsvoll sein, sondern allenfalls verantwortungsvoll genutzt werden kann. Das Konzept dahinter geht aber über die Frage der Nutzung hinaus und schließt je nach Begriffsverständnis in der Regel auch zum Beispiel Fragen nach der Güte des Trainingsdatensatzes ein.

Welche Medien geben sich überhaupt Leitlinien?

In liberalen Demokratien sind Medien in ihrer Berichterstattung meist frei und unterliegen keiner Kontrolle ihrer internen Arbeitsprozesse durch staatliche Stellen. Eine Pflicht für Medienhäuser, sich Richtlinien zum Umgang mit KI zu geben, gibt es daher nicht – ebenso wenig wie Zentralstellen, in denen entsprechende Leitlinien hinterlegt werden müssten. Daraus folgt eine Einschränkung: Es ist nicht ohne Weiteres möglich, die Zahl der Redaktionen zu benennen, die sich bis zum Stichtag der vorliegenden Untersuchung Ende April 2023 eigene KI-Richtlinien gegeben haben. Daher können nur jene Dokumente untersucht werden, die von den Redaktionen entweder selbst öffentlich gemacht oder gezielt der Forschung zur Verfügung gestellt worden sind.

Sieben Richtlinien sind Gegenstand der Studie. Sechs von ihnen wurden von den jeweiligen Organisationen selbst veröffentlicht, nur die Richtlinien der niederländischen Nachrichtenagentur Algemeen Nederlands Persbureau, kurz ANP, fanden über die sozialen Netzwerke den Weg in die Öffentlichkeit (vgl. ANP 2023). Die Untersuchung umfasst Dokumente von Medienorganisationen aus fünf westlichen Ländern, unter ihnen jeweils zwei aus Deutschland und den Vereinigten Staaten und je eine aus Kanada, Großbritannien und den Niederlanden. Mit insgesamt vier von sieben Organisationen stellen Nachrichtenagenturen klar den Schwerpunkt der Untersuchung, eingeschlossen sind hier neben der ANP aus Den Haag noch die als dpa abgekürzte Deutsche Presse-Agentur (vgl. DPA 2023) mit Sitz in Hamburg; beide berichten vor allem aus dem jeweiligen Land. Hinzu kommen zwei international ausgerichtete Agenturen, die in New York City ansässige Associated Press (vgl. AP o.J.), kurz AP, sowie die international tätige Agentur Thomson Reuters (vgl. THOMSON REUTERS o.J.) mit Sitz in Toronto. Dazu treten die Dokumente von zwei öffentlich-rechtlichen Rundfunkanstalten, dies sind der Bayerische Rundfunk (BR) aus München (vgl. BR 2020) sowie die British Broadcasting Corporation (BBC) mit Sitz in London (vgl. BBC 2021). Schließlich hat sich auch das amerikanische Technikmagazin *Wired* Richtlinien gegeben; die Redaktion des vom Verlag Condé Nast herausgegebenen Hefts befindet sich in San Francisco. Die in die Untersuchung einbezogenen Organisationen arbeiten also auf der Grundlage von teilweise unterschiedlichen nationalen Regeln für KI, etwa in Bezug auf den Datenschutz.

Bei vier von sieben Dokumenten ist der Zeitpunkt der Erstveröffentlichung angegeben. Wie aus Tabelle 1 hervorgeht, ist in zwei Fällen das genaue Datum bekannt, in zwei anderen ist der Zeitrahmen auf einen beziehungsweise zwei Monate genau angegeben. Über die chronologische Abfolge der Richtlinien lassen sich darum nicht durchweg belastbare Aussagen treffen. Mögliche Zusammenhänge bestehen allenfalls innerhalb der Mediengattungen. So hat der BR seine Richtlinien etwa ein halbes Jahr früher veröffentlicht als die britische

BBC. Auffallend ist auch, dass die zwei Nachrichtenagenturen dpa und ANP binnen kurzer Zeit im Frühjahr 2023 erstmals eigene Richtlinien veröffentlicht haben – etwa ein halbes Jahr, nachdem der amerikanische Anbieter OpenAI sein Sprachmodell GPT-3 in angepasster Form als ChatGPT am 30. November 2022 freigeschaltet hat (vgl. OPENAI 2022). Viel spricht dafür, dass dieses Ereignis und der darauffolgende »Hype« (MENN 2023) die Entwicklung von redaktionellen Leitlinien befeuert haben könnten. Auswertungen der amerikanischen Suchmaschine Google zeigen, dass das weltweite Interesse an KI seit Anfang Dezember deutlich gestiegen ist (vgl. GOOGLETRENDS 2023).

Tabelle 1
Übersicht der untersuchten Richtlinien

| | Hauptsitz | Kategorie | Datum der Erstveröffentlichung von KI-Richtlinien |
|--|---|---------------------------------------|---|
| Algemeen Nederlands Persbureau (ANP) | Den Haag, Niederlande | Nachrichtenagentur | März/April 2023 |
| Associated Press (AP) | New York City, Vereinigte Staaten von Amerika | Nachrichtenagentur | Ohne Datum |
| Bayerischer Rundfunk (BR) | München, Deutschland | Öffentlich-rechtliche Rundfunkanstalt | 30. November 2020 |
| British Broadcasting Corporation (BBC) | London, Großbritannien | Öffentlich-rechtliche Rundfunkanstalt | Mai 2021 |
| Deutsche Presse-Agentur (dpa) | Hamburg, Deutschland | Nachrichtenagentur | 3. April 2023 |
| Thomson Reuters | Toronto, Kanada | Nachrichtenagentur | Ohne Datum |
| Wired | San Francisco, Vereinigte Staaten von Amerika | Technik-Magazin | Ohne Datum |

Quelle: Eigene Darstellung

Spezifität, Struktur und Verbindlichkeit

Die Richtlinien unterscheiden sich zunächst in formaler Hinsicht, also mit Blick auf Spezifität, Struktur und Verbindlichkeit. Spezifität beschreibt die Frage, ob eine KI-Richtlinie explizit als solche gekennzeichnet ist. Der Begriff der Richt- oder Leitlinie soll hier vergleichsweise weit gefasst werden: Alle Schriften, die an zentraler Stelle wie zum Beispiel einem Editorial, einem Blogbeitrag oder einem gesonderten Dokument auf einen grundlegenden Umgang der jeweiligen Medienorganisation mit KI abzielen, sollen darunter zusammengefasst werden – auch dann, wenn der jeweilige Beitrag nicht explizit als solcher beschrieben ist. Mit Struktur ist gemeint, welchen formalen Aufbau das Dokument aufweist. Manche Redaktionen gießen ihre Handreichungen in einen

journalistisch anmutenden Fließtext, andere strukturieren ihr Papier strikt in mehrere Kapitel oder Punkte. Die Verbindlichkeit regelt schließlich, in welchem Umfang die jeweils festgeschriebenen Regeln gelten sollen.

Die Spezifität der Richtlinien unterscheidet sich deutlich. Die amerikanische Nachrichtenagentur Associated Press präsentiert ihre Regeln auf einer Überblicksseite, in der sie ihre KI-Aktivitäten grundlegend vorstellt (vgl. AP o.J.). Handreichungen für den Umgang mit KI werden darum auch nicht verbindlich als solche bezeichnet, sondern unter dem Begriff »strategy around the technology« (AP o.J.) behandelt. Deutlich klarer formulieren die drei Nachrichtenagenturen den Anspruch, in ihren Dokumenten Leitplanken für den Umgang mit KI zu ziehen. Die ANP verweist in ihren Leitlinien zudem auf die Chefredaktion, um die Verbindlichkeit der Regularien zu unterstreichen (vgl. ANP 2023). Auch die BBC macht klar, worum es geht. Dabei verwendet sie nicht den Begriff der KI, sondern des maschinellen Lernens, kurz ML (vgl. BBC 2021). Am deutlichsten macht das amerikanische Magazin *Wired*, welche redaktionelle Bindungskraft Leserinnen und Leser von dem Dokument erwarten dürfen: »How WIRED Will Use Generative AI Tools« (WIRED 2023, Hervorhebung im Original). Anstatt nüchtern Grundsätze zu referieren, kleiden die Verantwortlichen des Magazins ihre Richtlinien in eine sprachlich kraftvolle Absichtserklärung. Einen ähnlichen Ansatz wählt die niederländische Nachrichtenagentur (vgl. ANP 2023).

Bei der Struktur der Dokumente zeigen sich ebenfalls Unterschiede. Allgemein gilt: Je stärker ein Dokument strukturiert ist, desto spezifischer können die einzelnen Elemente sein – und desto präziser kann eine Richtlinie auf konkrete Anwendungsfälle im redaktionellen Kontext abzielen. Allein die ANP formuliert ihre Regeln vergleichsweise unstrukturiert (vgl. ANP 2023). Allein die ANP formuliert ihre Regeln vergleichsweise unstrukturiert (vgl. ANP 2023). Die Handreichungen der AP sind hingegen zwar eher allgemein, gleichwohl bezieht die Agentur ihre Strategie auf vier Ebenen; drei bilden den journalistischen Prozess ab, bei der vierten geht es um die Zusammenarbeit mit anderen Akteuren (vgl. AP o.J.). Die Macher von *Wired* beziehen ihre Ausführungen auf zwei inhaltliche Ebenen, also Texte und Bilder, untergliedert nach fünf beziehungsweise drei Punkten (vgl. WIRED 2023). Zwei Nachrichtenagenturen haben sich dafür entschieden, ihre Richtlinien auf jeweils fünf Punkte herunterzubrechen (vgl. DPA 2023; THOMSON REUTERS o.J.) – dazu passt, dass die Richtlinien sich inhaltlich vor allem auf den Kernbereich der Nachrichtenproduktion beschränken. Die dpa macht die Wahl des Umfangs der Richtlinien in einem Begleittext selbst zum Thema: Die Richtlinien sollen ihr dabei helfen, »den Umgang mit KI zu begleiten, ohne sich in einem Wald aus Regeln zu verlieren« (DPA 2023). Der BR differenziert seine Leitplanken feiner aus und strukturiert sein Dokument in zehn Punkte (vgl. BAYERISCHER RUNDFUNK 2020). Am stärksten untergliedert die BBC ihre Richtlinien. Die Anstalt beschränkt sich nicht auf die Darstellung

allgemeiner Kriterien, sondern legt ihr Dokument als eine Handreichung für interdisziplinäre Projektteams mit 47 einzelnen Punkten an (vgl. BBC 2021).

Kaum eine Leitlinie trifft differenzierte Aussagen zur Verbindlichkeit – es ist also davon auszugehen, dass die jeweils ausformulierten Kriterien in vollem Umfang und zu jedem Zeitpunkt einer möglichen Innovationsentwicklung Geltungskraft haben sollen. Nur der BR verfolgt das Konzept einer abgestuften Verbindlichkeit. Es sieht vor, dass umso mehr Kriterien erfüllt sein müssen, je näher eine geplante KI-Anwendung an die Umsetzung rückt (vgl. BR 2020).

3. Die Kunst der Selbstbeschränkung: Redaktionelle Richtlinien zwischen Verboten und Möglichkeiten

KI verspricht Redaktionen nahezu unbegrenzte Möglichkeiten im Umgang mit Texten und Bildern. In der Forschung wurde früh erkannt, dass die neuen Möglichkeiten der Personalisierung von Journalismus mit gesellschaftlichen Gefahren einhergehen – und Verzerrungen der öffentlichen Diskurse tendenziell vermindert werden können, wenn Medienhäuser sich Richtlinien geben (vgl. MARCONI 2020: 46). Die Organisationen versuchen, diesen Bedenken mit ihren freiwillig formulierten Leitlinien Rechnung zu tragen.

Die sieben Dokumente treffen dabei auf unterschiedlichen Ebenen jeweils andere Festlegungen. Im Folgenden geht es darum, inwiefern die Regeln auf Unternehmensziele und journalistische Werte bezogen werden, ob und welche Anwendungsbereiche und Grenzen formuliert sind, inwiefern die Leitlinien Festlegungen zu ›Responsible AI‹ treffen, wie es um Fragen der Transparenz und der menschlichen Kontrolle bestellt ist, welche Anforderungen an journalistische Kollaboration zu stellen und inwiefern die Dokumente als veränderbare Regelwerke aufzufassen sind.

Unternehmensziele und journalistische Werte

Jeder Versuch, eine redaktionelle Leitlinie zu entwerfen, baut auf der Frage auf, von was sich die jeweiligen Autorinnen und Autoren überhaupt leiten lassen wollen. In vielen Fällen haben die Medienhäuser in ihren Leitlinien Wert auf die Definition von strategischen Zielen oder journalistischen Werten gelegt. Auf diese Weise geben KI-Richtlinien oftmals einen tiefen Einblick in das journalistische Selbstverständnis einer Redaktion oder Organisation.

Einige Richtlinien unterstreichen die Rolle von Vertrauen. Dies sei »one of our most important values« (THOMSON REUTERS O.J.), heißt es bei Thomson Reuters. Die nachfolgenden Prinzipien sollen daher dem Zweck dienen »to promote trustworthiness in our continuous design, development, and deployment

of AI« (THOMSON REUTERS o.J.). Die Verantwortlichen in New York haben die Bedeutung von Vertrauenswürdigkeit nicht ohne Grund unterstrichen. Die Agentur fühlt sich nach eigener Darstellung der Geltung der sogenannten »Trust Principles« verpflichtet – einem Satz von Regeln, die eine freie und verlässliche Berichterstattung der Agentur sicherstellen sollen (vgl. THOMSON REUTERS 2018). Einen ähnlichen Ansatz verfolgt die BBC, allerdings belässt sie es nicht beim Konzept der Vertrauenswürdigkeit. Die Werte der Anstalt stellt die Richtlinie wie folgt dar: »upholding trust, putting audiences at the heart of everything we do, celebrating diversity, delivering quality and value for money and boosting creativity« (BBC 2021: 5). Die Fundamente, auf welche die britische Rundfunkanstalt ihre KI-Richtlinien stellt, sind demnach deutlich breiter. Die Anforderung, angesichts der Gebührenfinanzierung einen möglichst hohen Wert bereitzustellen, dürfte nicht ohne Grund Eingang in das Dokument gefunden haben – vielmehr ist sie im Kontext der politischen Debatte um die zukünftige Finanzierung der Rundfunkanstalt zu lesen (vgl. WATERSON 2022). KI im redaktionellen Kontext soll demnach einen zentralen Beitrag zur staatsbürgerlichen Bildung des Publikums leisten: »We will also seek to broaden, rather than narrow, our audience’s horizons« (BBC 2021: 6), heißt es darin.

Dass der Einsatz von KI mit den Argumenten der Effizienz und der gesellschaftlichen Verantwortung begründet wird, ist kein Alleinstellungsmerkmal des öffentlich-rechtlichen Rundfunks in Großbritannien. Bereits ein halbes Jahr zuvor hat der BR seine KI-Leitlinien veröffentlicht. Das Konzept des »Mehrwerts« (BR 2020) steht dabei für die Verantwortlichen der Richtlinien im Mittelpunkt. Der BR nutze KI, »um unsere Arbeit effizienter zu machen und verantwortlich mit den Ressourcen umzugehen, die uns die Beitragszahler:innen anvertrauen«, heißt es (BR 2020). KI solle auch dazu verwendet werden, »um neue Inhalte zu generieren, neue Methoden für investigative Recherchen zu entwickeln und Angebote für unsere Nutzer:innen attraktiver zu machen« (BR 2020). Auch in Deutschland wird regelmäßig über die Zukunft des gebührenfinanzierten Rundfunks diskutiert.

Privatwirtschaftlich organisierte Akteure bringen den Einsatz von KI tendenziell mit anderen Zielen in Verbindung als öffentlich-rechtliche Rundfunkanstalten. Die Deutsche Presse-Agentur begründet ihr Engagement mit journalistischen Wettbewerbsvorteilen. KI werde der Agentur »helfen, unsere Arbeit besser und schneller zu machen – immer im Sinne unserer Kunden und unserer Produkte« (DPA 2023), stellt sie klar. Etwas zurückhaltender formuliert es die ANP aus Den Haag. Sie betont ihr »Bekenntnis zu Qualität und Zuverlässigkeit« (ANP 2023)²¹ und verweist auf die Regelungen des Redaktionsstatus. Darin wird ausgeschlossen, dass der Journalismus der Agentur von außen beeinflusst wird,

2 Dieses und die folgenden Zitate aus dem Niederländischen wurden mithilfe von DeepL übersetzt.

außerdem betont sie die Werte der Unparteilichkeit und der Sorgfalt (vgl. ANP 2023).

Anwendungsbereiche und Grenzen von KI

Etwa die Hälfte der Medienorganisationen, die sich Leitlinien für den Umgang mit KI gegeben haben, umreißen darin die möglichen Anwendungsbereiche der neuen Technik. Einerseits definieren Redaktionen so jene Bereiche des journalistischen Prozesses, in denen ihnen der Einsatz von KI lohnenswert erscheint – sie erstellen also gewissermaßen eine Positivliste für KI im Newsroom. Auf der anderen Seite nutzen Unternehmen die Richtlinien, um eine Negativliste von Anwendungen zu bestimmen, die aus ihrer Sicht keinesfalls durch Algorithmen in ihrer journalistischen Integrität beeinträchtigt werden sollten.

Der BR und die Deutsche Presse-Agentur beschränken sich in ihren Richtlinien auf abstrakte Prinzipien. Nur unwesentlich klarer umreißt die Associated Press die möglichen Anwendungsfelder, dies sind Themenfindung (»to break news and dig deeper«, AP o.J.), Produktion (»to streamline workflows«, AP o.J.) und Distribution. Die genannten Beispiele bleiben aber vergleichsweise unspezifisch und entsprechen dem journalistischen Common Sense bei der Anwendung von KI (vgl. BECKETT 2019).

Wired beschränkt sich in seinen durch und durch journalistisch aufgeschriebenen Richtlinien hingegen nicht darauf, mögliche Anwendungsfelder im Rahmen einer Positivliste zu benennen. Vielmehr differenziert die Redaktion für den Umgang mit Texten wie auch mit Bildern, inwiefern sie KI überhaupt nicht oder möglicherweise nutzen will – oder damit vielleicht einmal experimentiert. *Wired* stellt klar, dass das Heft keine Geschichten veröffentlicht, deren Text ganz oder in Teilen von KI generiert wurde (vgl. WIRED o.J.). Grund sind die Grenzen der aktuellen Textgeneratoren, die sich mit dem journalistischen Anspruch der Redaktion sowohl in inhaltlicher als auch in stilistischer Weise offenbar nicht vertragen: »The current AI tools are prone to both errors and bias, and often produce dull, unoriginal writing« (WIRED o.J.). Die Redaktion fühlt sich schließlich gar in ihrer journalistischen Ehre gekränkt, sollte ein von ihr veröffentlichter Text auf der Grundlage eines Algorithmus' erstellt worden sein: »we think someone who writes for a living needs to constantly be thinking about the best way to express complex ideas in their own words« (WIRED o.J.). Aus diesen Gründen lehnt das Magazin es auch ab, dass journalistische Texte von einer KI redigiert werden. Denkbare Anwendungsbereiche beschränken sich stattdessen auf Versuche, KI Überschriften oder kurze Texte für soziale Netzwerke texten zu lassen; darüber hinaus soll KI auch genutzt werden können, um Ideen für mögliche Themen zu gewinnen. Auch bei der eigentlichen Recherche sollen Werkzeuge wie ChatGPT der Redaktion grundsätzlich behilflich sein können,

wenn auch unter Auflagen. So könnten Sprachmodelle beispielsweise genutzt werden, um sich durch große Dokumentenmengen zu lesen – sie würden dann genutzt wie die Google-Suchmaschine oder Wikipedia. Analog will *Wired* auch mit Bildern verfahren, die von KI – wie etwa dem von OpenAI entwickelten Werkzeug Dall-E – produziert worden sind.

Noch etwas weiter geht die ANP aus Den Haag. Die Journalisten sehen »viele Möglichkeiten« (ANP 2023), KI zu verwenden und sich von Algorithmen inspirieren zu lassen – etwa beim Schreiben von Überschriften und Hintergründen. Klare Grenzen für die Verwendung von Textgeneratoren definieren die Autoren nicht. Der Einsatz von Künstlicher Intelligenz im Newsroom bleibt demnach »den Redakteuren überlassen« (ANP 2023).

Merkmale von ›Responsible AI‹

Je stärker Instrumente der KI genutzt werden, desto mehr wächst der Einfluss dieser Systeme auf das Leben der Menschen. Damit die steigende Wirkung von Algorithmen nicht zu »Abhängigkeiten [...] oder Anpassungsdruck« (DEUTSCHER ETHIKRAT 2023) führt, gilt es, KI in einer sozial verantwortungsvollen Weise einzusetzen. Gemeinhin wird sogenannte ›Responsible AI‹ auf drei inhaltliche Dimensionen bezogen: Zurechenbarkeit, Verantwortung und Transparenz (vgl. DIGNUM 2019: 52f.). Dass der Mensch dabei im Mittelpunkt stehen soll, wie es in der Bundesrepublik unlängst der Deutsche Ethikrat in seiner Stellungnahme zur KI propagiert hat (vgl. DEUTSCHER ETHIKRAT 2023), kann als internationaler Konsens des globalen Westens gelten. Informatiker weisen darauf hin, dass ein gutes technisches Design von KI-Anwendungen es grundsätzlich möglich macht, ein hohes Maß an menschlicher Kontrolle mit einem hohen Grad an angestrebter Automatisierung zu verknüpfen (vgl. SHNEIDERMAN 2022: 79). Vor dem Hintergrund der Debatte um Kriterien und Umsetzung von ›Responsible AI‹ haben einige Medien ihre Leitlinien um Aspekte ergänzt, die sich auf Datenschutz sowie die Güte von Algorithmen, Trainings- und sonstigen Daten beziehen.

Beim Datenschutz setzen die jeweils geltenden Gesetze den Rahmen für KI-Anwendungen aller Art, so bezieht sich unter anderem die ANP direkt darauf (vgl. ANP 2023). Thomson Reuters gibt nur an, sie wolle »prioritize safety, security, and privacy throughout the design, development and deployment of our AI products and services« (THOMSON REUTERS O.J.). In Deutschland, wo die europäische Datenschutzgrundverordnung einen Großteil der rechtlichen Leitplanken zieht, betont der BR das Konzept der Datensparsamkeit. Man erhebe »so wenige Daten wie möglich und so viele wie nötig, um unseren Auftrag zu erfüllen« (BR 2020). Die britische BBC verspricht zunächst recht allgemein: »we will ensure that data is handled securely« (BBC 2021: 6). Allerdings spannt sie in den darauffolgenden Leitfragen für KI-Projektteams etliche Punkte auf, die ins Detail

gehen. Was die Nutzung von Daten angeht, so verweisen die Prinzipien auf ein anderes Dokument der Anstalt, die zuletzt im Jahr 2023 aktualisierte »privacy promise« (BBC 2023). Sie regelt Fragen der Transparenz, der Auswahlmöglichkeiten für das Publikum und der Datennutzung durch die BBC. Darüber hinaus sollen Projektteams alle Datennutzungen und -veränderungen dokumentieren, mögliche Fehler in den Datensätzen korrigieren, sowie die rechtliche Zulässigkeit der genutzten Daten prüfen, einschließlich ihrer Übereinkunft mit den Grundsätzen der Datenschutzgrundverordnung der Europäischen Union (vgl. BBC 2021: 15f.).

Was die Güte der Algorithmen angeht, so finden sich in den Leitlinien viele allgemeine Festlegungen. Die Deutsche Presse-Agentur etwa verspricht, nur »rechtmäßige KI« (DPA 2023) einzusetzen, »die sich an geltendes Recht und gesetzliche Bestimmungen hält und die unseren ethischen Grundsätzen gerecht wird, also etwa der menschlichen Autonomie, Fairness und demokratischen Werten« (DPA 2023). Thomson Reuters fasst die Grundsätze noch allgemeiner und formuliert sie als Absichtserklärung – die Agentur strebt demnach einen menschenzentrierten Ansatz an. Man wolle Werkzeuge entwickeln und benutzen »that treat people fairly« (THOMSON REUTERS O.J.). Einen ähnlichen Anspruch verfolgt die BBC, die von ihren Algorithmen verlangt, sie sollten »serve our audiences equally & fairly« (BBC 2021: 6).

In der Praxis beschreibt das Kriterium der Fairness einen Algorithmus, der mithilfe von ausgewogenen Trainingsdaten geschult worden ist und auf diese Weise kaum oder nur wenig verzerrte Ergebnisse produziert. Der BR und die ANP gehen in ihren Leitlinien auf das Problem der algorithmischen Verzerrung ein. Die Agentur beschränkt sich darauf, das Risiko zu benennen und dies bei der Bewertung von Werkzeugen zu berücksichtigen (vgl. ANP 2023). Allgemein setzt sie auf das Kriterium der »Zuverlässigkeit« (ANP 2023). Differenzierter behandelt der BR das Thema. Von Dienstleistern verlange man demnach »belastbare Informationen zu den Datenquellen« (BR 2020) und diskutiere grundsätzlich, auch bei internen Entwicklungen, die »Integrität und Qualität der Trainingsdaten« (BR 2020). Eine Verringerung der algorithmischen Verzerrung helfe dabei, die von der Rundfunkanstalt in ihren Leitlinien herausgehobene »gesellschaftliche Vielfalt abzubilden« (BR 2020). Auch die BBC weist ihre Projektteams darauf hin, die zugrundeliegenden Trainingsdaten auf möglichen »bias« (BBC 2021: 19) hin zu überprüfen und diesen bei Bedarf zu korrigieren. Einzig der BR geht in seinen Richtlinien nicht nur auf die Güte der Trainingsdaten ein, sondern auch auf die Qualität der übrigen Daten, mit denen das Modell arbeitet. In München will man sich dazu verpflichten, bei den Mitarbeiterinnen und Mitarbeitern das »Bewusstsein für den Wert von Daten und die konsequente Pflege von Metadaten« (BR 2020) zu stärken, da belastbare Anwendungen der KI nur mit »verlässlichen Daten« (BR 2020) entwickelt werden könnten.

Transparenz beim journalistischen Endprodukt

Wann immer KI an einem journalistischen Text in maßgeblicher Weise mitgewirkt hat, stellt sich eine entscheidende Frage: »Who should get the byline?« (MARCONI 2020: 97). Während es bislang unüblich war, Sprachmodelle wie ChatGPT und Bard in die Autorenzeile aufzunehmen, so haben etliche Redaktionen doch zumindest in unterschiedlicher Weise damit begonnen, den Beitrag von KI zum Entstehen eines journalistischen Werkstücks kenntlich zu machen – so zum Beispiel neben AP auch die britische Tageszeitung *The Guardian* und das amerikanische *Wall Street Journal* (vgl. MARCONI 2020: 97f.). Es überrascht darum nicht, dass die Frage der Transparenz in den Richtlinien vielfach thematisiert wird – und denn auch durchweg zugunsten einer möglichst ausgeprägten Sichtbarkeit. Allerdings machen die Richtlinien oftmals keine Angaben, wie genau der Beitrag von KI gekennzeichnet werden soll.

Drei Viertel der untersuchten Nachrichtenagenturen verpflichten sich dazu, die Verwendung von KI zu kennzeichnen. Thomson Reuters verspricht, man wolle dessen Nutzung »explainable« (THOMSON REUTERS o.J.) machen. In ähnliche Worte kleidet die dpa ihre Selbstverpflichtung: »Dort, wo Inhalte ausschließlich durch KI erzeugt werden, machen wir dies transparent und erklärbar« (DPA 2023). Und auch die niederländische ANP verspricht, gegenüber ihren Kunden »offen über unsere Arbeitsweise und den Einsatz von KI oder anderen technischen Systemen« (ANP 2023) zu sein. Das Versprechen, man bemühe sich in Den Haag »um Transparenz« (ANP 2023), wird an anderer Stelle allerdings wieder abgeschwächt: Man erwähne »dort, wo wir es als Redaktion für angebracht halten, den Umfang der KI-Nutzung« (ANP 2023).

Das Magazin *Wired* verpflichtet seine Autoren, den Beitrag von KI zur Entstehung des jeweiligen Beitrags kenntlich zu machen – andernfalls wolle man dies als gleichwertig zu einem Plagiat werten (vgl. WIRED o.J.). Das Magazin beabsichtigt auch, wo es möglich ist, die Quellen der KI offenzulegen. Beim Bayerischen Rundfunk will man transparent machen, »welche Technologien wir einsetzen, welche Daten wir erheben und welche Redaktionen oder Partner die Verantwortung dafür tragen« (BR 2020). Und die Verantwortlichen in München gehen noch einen Schritt weiter. Sollten Probleme im Umgang mit KI auftreten, wolle man das »zum Thema selbstreflexiver Berichterstattung« (BR 2020) machen. Auch wenn die britische BBC in ihren Handreichungen nicht so weit geht, zeichnet sie sich in Sachen Transparenz an anderer Stelle aus. Die Verantwortlichen in London leiten KI-Teams an, ihr Projekt in ein eigenes Register für interne KI-Anwendungen einzutragen (vgl. BBC 2021: 21). Darüber hinaus soll die Funktionsweise der Anwendung nicht nur den Beschäftigten der britischen Anstalt verdeutlicht, sondern auch dem Publikum »in plain English« (BBC 2021: 21), also unumwunden und ohne Fachsprache, erläutert werden.

Bemerkenswert ist, dass die Frage, ob KI eine Erwähnung in der Autorenzeile eines Beitrags rechtfertigt, zuletzt weniger vom Journalismus thematisiert wurde als von der Wissenschaft. Anfang des Jahres haben die Fachzeitschriften *Nature* und *Proceedings of the National Academy of Sciences of the United States of America* (PNAS), in Editorials klargestellt, dass sich Sprachmodelle wie ChatGPT nicht für eine Nennung als Autor qualifizieren, da diese nicht für die Ergebnisse zur Rechenschaft gezogen werden können (vgl. NATURE 2023; PNAS 2023).

Kontrolle durch Menschen

Die Frage, inwiefern der Einsatz von KI bei der Produktion von journalistischen Inhalten unter menschlicher Kontrolle stehen soll, ist derzeit im Mittelpunkt vieler Überlegungen. KI hat es Redaktionen früh ermöglicht, auf der Grundlage strukturierter Daten automatisierte Texte zu erstellen (vgl. DIAKOPOULOS 2019: 96f.). Diese Möglichkeiten wurden durch große Sprachmodelle massiv erweitert – zumindest hypothetisch, denn in der Praxis machen die teilweise inhaltlich fehlerhaften Texte der Generatoren oft Probleme. Nicht nur in der Journalismusforschung ist daher auf die Bedeutung menschlicher Kontrolle bei der Nutzung von Textbausteinen aus der Feder von KI hingewiesen worden (vgl. MARCONI 2020: 49), auch Redaktionen setzen sich mit dieser Frage intensiv auseinander (vgl. WOLFENAGEL 2023).

Es verwundert daher nicht, dass alle sieben Richtlinien die Frage der menschlichen Kontrolle von Beiträgen der KI zu finalen journalistischen Produkten adressieren – wenn auch in unterschiedlichem Umfang und mit wechselnder inhaltlicher Einordnung. *Wired* ist hierbei abermals am restriktivsten. Die Richtlinien des Magazins sehen vor, dass keinerlei Texte veröffentlicht werden, an denen KI mitgeschrieben hat oder die von ihr redigiert worden sind (vgl. WIRED o.J.). Mit seiner vorsichtigen Haltung nimmt das amerikanische Magazin eine Außenseiterrolle ein. Fünf von sieben Richtlinien regeln, dass Menschen unter bestimmten Umständen eine redaktionelle Kontrolle jener Textbausteine vornehmen, die von KI erstellt worden sind. Die Abnahme der jeweiligen Inhalte ist mal zwingend vorgeschrieben, ein anderes Mal nur unter bestimmten Bedingungen.

Die Richtlinien der beiden Nachrichtenagenturen dpa und ANP sehen vor, dass ein Mensch alle journalistischen Inhalte vor Veröffentlichung abnimmt, die von oder mit der Unterstützung von KI erstellt worden sind. Bei der Deutschen Presse-Agentur heißt es: »Die dpa setzt KI nur unter menschlicher Aufsicht ein« (DPA 2023). Die Redaktion betont, dass ein Mensch die »letzte Entscheidung« (DPA 2023) über den Einsatz KI trifft. Für eine ähnliche Konditionierung hat sich die ANP entschieden: »Wir können KI oder ähnliche Systeme zur Unterstützung der Endredaktion einsetzen, vorausgesetzt, ein Mensch nimmt anschließend

eine Endkontrolle vor« (ANP 2023). Weiter spezifizieren die Verantwortlichen, dass KI vor allem bei redaktionellen Zwischenschritten eingesetzt werden kann: »In unserer Produktionskette halten wir uns an die derzeit gültige Linie Mensch-Maschine-Mensch« (ANP 2023). Von der KI generierte Inhalte würden nicht verwendet, »ohne diese Informationen durch einen Menschen zu überprüfen« (ANP 2023).

Etwas weniger strikt sind die Nachrichtenagentur Thomson Reuters sowie die beiden untersuchten Rundfunkanstalten. Ihre jeweiligen Richtlinien sehen eine verbindliche Kontrolle der journalistischen Endprodukte nur unter bestimmten Bedingungen vor, aber anders als die beiden Agenturen aus Deutschland und den Niederlanden eben nicht in jedem Fall. Thomson Reuters formuliert die Rolle des Menschen nicht nur weniger verbindlich, sondern kleidet sie auch in die sprachlich schwächere Form einer Absichtserklärung: »Thomson Reuters will strive to maintain a human-centric approach«, (THOMSON REUTERS o.J.) heißt es. Wie dieser Ansatz genau beschaffen ist, geht aus dem Dokument nicht hervor. Gleiches gilt für die vergleichsweise vage formulierte Richtlinie, dass die Organisation für die Produkte und Dienstleistungen, bei denen KI verwendet wird, die Verantwortung trägt: »Thomson Reuters will maintain appropriate accountability measures for our AI products and services« (THOMSON REUTERS o.J.).

Unter dem Schlagwort »Human in the loop« (BBC 2021: 6) beschreibt die BBC die Rolle von Redakteurinnen und Redakteuren. Eine eindeutige Festlegung, in welchem Bereich der Mensch die Arbeit von KI kontrollieren und abnehmen muss, geht damit allerdings nicht einher. Es handele sich um eine fortlaufende Entwicklung, die BBC wolle damit experimentieren. »Algorithms form only part of the content discovery process for our audiences, and sit alongside (human) editorial curation« (BBC 2021: 6). Die Einlassungen lassen sich als eine grundlegende Einbeziehung der Redaktion bei der Nutzung von KI lesen, die allerdings weiterer Präzisierung bedarf.

Dem BR gelingt es, eine differenzierte und gleichzeitig klare Haltung bei der Rolle menschlicher Kontrolle auszuformulieren. Die Rundfunkanstalt setzt redaktionelle KI-Inhalte zunächst unter einen allgemeinen Abnahmevorbehalt, ergänzt um eine dynamische Öffnungsklausel. »Auch bei *automated journalism* und Datenjournalismus liegt die publizistische Verantwortung bei den Redaktionen. So bleibt das Abnahmeprinzip bei automatisiert erstellten Inhalten bestehen«, stellen die Verantwortlichen klar (BR 2020, Hervorhebung im Original). »Es entwickelt sich allerdings weiter: Das Prinzip der Einzelprüfung wird zu einer Plausibilitätsprüfung von kausalen Zusammenhängen in der Datenstruktur und einem rigorosen Integritätstest der Datenquelle« (BR 2020). Anstatt jeden einzelnen Beitrag, an dessen Genese KI beteiligt war, einzeln redaktionell abzunehmen, öffnet der BR die Tür für eine ausschließliche redaktionelle Kontrolle der technischen Funktionsweise des entsprechenden KI-Instruments.

Anforderungen an die redaktionelle Zusammenarbeit

Wenn KI im Newsroom zum Einsatz kommt, ist das meist das Ergebnis einer Zusammenarbeit unterschiedlicher Fachabteilungen: Journalisten und Informatikerinnen, Datenspezialisten und Produktmanagerinnen arbeiten Hand in Hand – und nicht selten auch mit externen Dienstleistern, da nicht jede Redaktion über alle Kompetenzen verfügt, die für den Roll-Out einer KI-Anwendung erforderlich sind. Daher regeln einige Organisationen in ihren Richtlinien zum Umgang mit KI neben journalistischen Fragestellungen auch Aspekte der interdisziplinären Zusammenarbeit sowohl im Innen- wie auch im Außenverhältnis.

Was die Rollenverteilung innerhalb der Organisation betrifft, ist keine der untersuchten Leitlinien so detailliert wie das Dokument der BBC. Es benennt einerseits unterschiedliche Rollen, die jeweils involviert sein können, unter anderem Produktmanager sowie Beschäftigte der Fachabteilung für Qualität, Risiko und Versicherung (vgl. BBC 2021: 21). Vor allem aber richtet sich das Dokument mit seinen ausführlichen Leitfragen, die klar auf technische und dokumentarische Prozesse abzielen, auf Mitarbeiter jenseits des Newsrooms (vgl. BBC 2021: 17f.). Weniger deutlich wird in diesem Punkt der BR, der klarstellt, dass KI-Projekte durch »möglichst diverse Teams« (BR 2020) ermöglicht werden sollten. Ähnlich unspezifisch stellt die dpa ihre Anforderungen an die Zusammenarbeit dar. In den Richtlinien heißt es, man ermuntere »alle Mitarbeiterinnen und Mitarbeiter« (DPA 2023), sich dem Thema KI gegenüber zu öffnen – womit zwar vor allem, aber gewiss nicht ausschließlich Reporterinnen oder Redakteure gemeint sein dürften.

Dass Medienorganisationen über ihre Leitlinien regeln, inwiefern sie bei KI-Projekten mit Akteuren außerhalb des eigenen Hauses zusammenarbeiten, ist die Ausnahme. Die AP berichtet, dass sie mit Start-Ups zusammenarbeite, um von »external innovation« (AP o.J.) zu vergleichsweise geringen Kosten profitieren zu können. Darüber hinaus schmiedete man in New York auch Partnerschaften mit weiteren Institutionen, unter anderem mit den Investmentunternehmen Social Starts und Matter Ventures sowie dem NYC Media Lab, einer Zusammenarbeit zwischen mehreren Universitäten und Unternehmen der Stadt (vgl. AP o.J.). Die Zusammenarbeit mit Hochschulen und Universitäten ist auch Gegenstand der Leitlinien des Bayerischen Rundfunks. Der »Austausch mit wissenschaftlichen Instituten und KI-Ethik-Expert:innen« (BR 2020) soll die interdisziplinäre Herangehensweise an das Thema KI prägen.

Dynamik der Richtlinien

Mit dem Aufkommen neuer Textgeneratoren und weiterer Werkzeuge verändern sich die Anwendungsmöglichkeiten von KI im redaktionellen Kontext in hoher

Geschwindigkeit. Um auch auf absehbare Zeit eine angemessene Orientierung zu ermöglichen, können Medienorganisationen ihre Leitlinien inhaltlich möglichst weit fassen, wie es unter anderem Thomson Reuters und der BR tun. Die andere – nicht selten auch zusätzlich ins Werk gesetzte – Möglichkeit ist, die Richtlinien gezielt als vorübergehendes Regelwerk zu beschreiben.

Mehrere Redaktionen haben eine dynamische Komponente in ihre KI-Regeln integriert. *Wired* spricht den weiteren Wandel von KI an und stellt in Aussicht, dass sie »may modify our perspective over time« (WIRED o.J.). Mögliche Änderungen würden in dem Dokument transparent gemacht, heißt es weiter. Auch die Nachrichtenagenturen sind sich bewusst, dass die aktuellen Regeln zum Umgang mit KI kaum in Stein gemeißelt sein dürften. »These AI principles will evolve as the field of AI and its applications matures« (THOMSON REUTERS o.J.) schreibt Thomson Reuters. Ähnlich stellt die ANP klar, dass es sich bei ihren Richtlinien um ein »lebendes Dokument« (ANP 2023) handele, »das von den Chefredakteuren angepasst werden kann, wenn die Entwicklungen dies erfordern« (ANP 2023).

Der BR benennt die Anforderung, die Richtlinien anzupassen, im Zusammenhang mit einer allgemeinen Dynamik von journalistischen KI-Anwendungen. »Experimente sind Teil des Prozesses«, schreiben die Autorinnen und Autoren (BR 2020). Zur möglichen Anpassung der Leitlinien passt, dass unter allen untersuchten Richtlinien allein der BR ein abgestuftes Modell der Verbindlichkeit entworfen hat; eine KI-Anwendung muss demnach umso mehr – und schließlich alle – Kriterien der Richtlinien erfüllen, je näher sie an der Veröffentlichung steht. Dabei berücksichtigt die Richtlinie auch, dass es bei der Anwendung von KI zu »ethischen Grenzsituationen« (BR 2020) kommen kann. »Die Erfahrungen evaluieren wir aus der Perspektive des Medienstaatsvertrags und der hier fixierten Richtlinien« (BR 2020).

Dass sich der Einsatz von KI mit der Zeit verändert, ist auch in den Richtlinien der BBC zum maschinellen Lernen Thema. »ML is an evolving set of technologies, where the BBC continues to innovate and experiment«, (BBC 2021: 6) heißt es dort. Es besteht die Möglichkeit, die Fragestellungen im Rahmen der BBC-Checkliste zu überarbeiten. Eine der Leitfragen lautet: »What important changes (or re-visioning / redeployment of the model) would trigger a MLEP checklist review?« (BBC 2021: 26).

4. Diskussion der Ergebnisse: Jeder reguliert für sich allein

Wie nähert man sich dem Neuen, dem Unbekannten? Die Untersuchung der Frage, mit welchen Richtlinien sich internationale Medienorganisationen erste Regeln für den Umgang mit KI gegeben haben, zeigt: Jede Redaktion entscheidet

das zunächst für sich allein. Denn die Leitlinien unterscheiden sich erheblich voneinander, sowohl formal als auch inhaltlich. So etwas wie ein Standardmodell der Richtlinien hat sich bis Ende April 2023 noch nicht herausgebildet.

Welche KI-Regeln sich internationale Medien geben

In der Form haben die untersuchten Organisationen – unter ihnen ein Magazin, vier Nachrichtenagenturen und zwei Rundfunkanstalten – vor allem auf kurze, nüchtern formulierte Leitlinien gesetzt, in denen sie knapp in einer definierten Anzahl von Punkten darlegen, wie sie mit KI umgehen wollen. Vereinzelt wurden die Richtlinien auch in Form eines sprachlich an den Magazinjournalismus angelehnten Textes gegossen oder im Rahmen eines Werkzeugs entfaltet, das interdisziplinäre Projektteams mit Leitfragen auf wichtige Punkte zum Umgang mit Algorithmen hinweist. Dabei zeigten sich grobe Zusammenhänge zwischen dem journalistischen Stil der jeweiligen Organisation und der gewählten Form der Richtlinie: Die Nachrichtenagenturen präsentierten ihre Leitlinien eher kurz und nachrichtlich, das Magazin verfolgte einen erzählerischen Stil und eine für ihre strukturelle Komplexität bekannte britische Rundfunkanstalt wählte die Form eines ausführlichen Leitfadens. Fast alle Richtlinien waren als solche zu erkennen, nur in einem Fall wurde die Herangehensweise in die allgemeine Darstellung eigener KI-Aktivitäten eingebettet. Auch verzichteten sechs der sieben Organisationen darauf, Abstufungen bei der Verbindlichkeit der Regeln vorzunehmen; nur eine deutsche Rundfunkanstalt setzte den Erfüllungsgrad der Richtlinien in Bezug zum Entwicklungsfortschritt eines KI-Projekts. Insgesamt wurden jeweils zwei Medienorganisationen aus Deutschland und den Vereinigten Staaten sowie je eine aus den Niederlanden, Großbritannien und Kanada in die Untersuchung eingeschlossen.

Auf der inhaltlichen Ebene zeigt der Vergleich der sieben internationalen Leitlinien deutliche Unterschiede. Die von den Redaktionen bearbeiteten Inhalte lassen sich grob in sieben Kategorien unterteilen. Die nachfolgende Tabelle 2 zeigt in der Übersicht, welche Medien in ihren Leitlinien zu welchen Aspekten Festlegungen gleich welcher Art getroffen haben.

Im arithmetischen Mittel haben die untersuchten Organisationen in ihren KI-Richtlinien 4,9 von sieben inhaltlichen Dimensionen berücksichtigt. Am breitesten fielen die Leitlinien des Bayerischen Rundfunks aus, die alle sieben Punkte umfassen, gefolgt von den Dokumenten der Nachrichtenagenturen Algemeen Nederlands Persbureau und Deutsche Presse-Agentur mit jeweils sechs Dimensionen. Auch die britische Rundfunkanstalt BBC erreicht diesen Wert. Mit jeweils vier inhaltlichen Punkten liegt die Regelungsbreite der kanadischen Nachrichtenagentur Thomson Reuters und des amerikanischen Magazins *Wired* leicht unter dem Mittel. Mit nur einer von sieben Dimensionen hat sich

Tabelle 2

Inhaltlicher Regelungsumfang von KI-Richtlinien internationaler Medien

| | Journalistische Ziele und Werte | Anwendungsbereiche | »Responsible AI« | Regeln für Transparenz | Menschliche Kontrolle | Formen der Zusammenarbeit | Dynamik der Regeln |
|-----------------|---------------------------------|--------------------|------------------|------------------------|-----------------------|---------------------------|--------------------|
| ANP | + | + | + | + | + | - | + |
| AP | - | - | - | - | - | + | - |
| BR | + | + | + | + | + | + | + |
| BBC | + | - | + | + | + | + | + |
| dpa | + | + | + | + | + | + | - |
| Thomson Reuters | + | - | + | + | + | - | - |
| <i>Wired</i> | - | + | - | + | + | - | + |

Quelle: Eigene Darstellung

die amerikanische Nachrichtenagentur Associated Press für einen vergleichsweise geringen Umfang der Selbstregulierung entschieden. Das dürfte mit dem besonderen Charakter der Leitlinien zusammenhängen, die als einzige nicht explizit als solche beschrieben sind.

Betrachtet man die einzelnen Bereiche genauer, fallen schnell inhaltliche Schwerpunkte auf. Im arithmetischen Mittel wird jede einzelne Sphäre der Selbstregulierung von 4,9 Organisationen in ihren Richtlinien thematisiert. Gleichwohl treffen fast alle untersuchten Redaktionen – es sind sechs von sieben – Aussagen zur Rolle der menschlichen Kontrolle, wenn KI für journalistische Endprodukte genutzt wird. Das lässt darauf schließen, dass die Frage, ob und wann eine Journalistin oder ein Journalist den Beitrag des Computers überprüfen oder redaktionell abnehmen soll, die meisten Medienhäuser derzeit umtreibt. Genauso oft treffen die Leitlinien auch Aussagen zur Frage, inwiefern der Beitrag von KI für das Publikum transparent gemacht werden soll. Auf strategische Unternehmensziele beziehungsweise journalistische Werte sowie Anforderungen an vertrauenswürdige KI [»Trustworthy AI«] haben fünf von sieben Organisationen hingewiesen, um ihre KI-Aktivitäten zu begründen. Genauso viele stellten Anforderungen an vertrauenswürdige KI. Seltener ging es in den untersuchten Leitlinien um die Definition von Anwendungsbereichen, über Formen der interdisziplinären Zusammenarbeit sowie um eine mögliche Aktualisierung der Leitlinien. In allen diesen Fällen befassten sich in wechselnden Zusammensetzungen jeweils nur vier von sieben Redaktionen damit.

Dass der Mensch auch in Zeiten des KI-unterstützten Journalismus' im Mittelpunkt stehen soll, stellt zusammengefasst den Konsens der untersuchten Leitlinien dar. Von den sechs Richtlinien mit dezidierten Festlegungen bleiben BBC und Thomson Reuters in ihren Ausführungen vage. Die verbleibenden vier Richtlinien schreiben den Redakteurinnen und Redakteuren demgegenüber eine vergleichsweise starke Position zu. *Wired* verbietet den Einsatz von KI-Textbausteinen beziehungsweise einem KI-unterstützten Redigat pauschal, zudem sehen die Leitlinien der europäischen Nachrichtenagenturen dpa und ANP keine Ausnahmen von der Pflicht vor, dass ein Mensch die entsprechenden Texte kontrolliert. Lediglich der BR sieht eine Öffnung des Kontrollprinzips vor, wenn dieses durch die strikte Überprüfung der Arbeitsweise der Algorithmen ersetzt wird. Damit schieben die meisten Organisationen, die in die Untersuchung eingeschlossen wurden, einer unreflektierten und unkritischen Übernahme von Textmaterial, das von KI erstellt oder maßgeblich geprägt wurde, einen Riegel vor.

Drei von vier Nachrichtenagenturen verpflichten sich dazu, die Verwendung von KI zu kennzeichnen und damit den Nutzerinnen und Nutzern transparent zu machen, dazu gehören Thomson Reuters, dpa und ANP. Auch *Wired*, das insgesamt einen restriktiven Umgang mit KI vertritt, trifft dazu entsprechende Regelungen. Für Transparenz treten auch die beiden untersuchten Rundfunkanstalten aus Europa, BR und BBC, ein. Damit wird deutlich, dass das Kriterium der Transparenz zumindest bei den untersuchten Medien fast durchweg unumstritten ist.

Etwa die Hälfte der Organisationen befassen sich in ihren Richtlinien mit möglichen Anwendungsbereichen von KI. Der BR und die dpa bleiben in ihren Darstellungen vergleichsweise abstrakt, die AP nennt einzelne Beispiele entlang des journalistischen Produktionsprozesses. Ausdrücklich auf die Nutzung von KI-generierten Bildern geht nur *Wired* ein; ähnlich wie bei Texten positioniert sich das amerikanische Magazin mit Blick auf eine Veröffentlichung der Ergebnisse ablehnend.

Bemerkenswert ist, dass nur fünf von sieben Redaktionen sich in ihren Leitlinien mit der Frage beschäftigen, welche Anforderungen an ›Responsible AI‹ gerichtet werden müssen und was das für die redaktionelle Nutzung der neuen Techniken bedeutet. Der Frage kommt vor dem Hintergrund einer in der Forschung bereits thematisierten möglichen Abhängigkeit der Medienorganisationen von den Anbietern leistungsstarker KI-Engines (vgl. SIMON 2022) ein besonderes Gewicht zu. Die wichtige inhaltliche Dimension lässt sich weiter in einzelne Elemente zerlegen. Anforderungen an den Datenschutz formulieren vier Medien, dies sind ANP, Thomson Reuters, der BR sowie die BBC. Die Güte von Algorithmen ist ausdrücklich für die dpa und abermals die BBC, den BR und Thomson Reuters ein Thema. Auf Probleme, die aus algorithmischer Verzerrung resultieren, weisen nur die beiden Rundfunkanstalten hin.

Ausblick

Die Untersuchung von sieben Richtlinien für den Umgang mit KI zeigt, dass internationale Medienorganisationen sich intensiv mit wesentlichen Fragen der neuen Technik befassen. Während menschliche Kontrolle und Transparenz gegenüber dem Publikum die Autorinnen und Autoren der Dokumente am meisten beschäftigt haben, zeigt die inhaltliche Auswertung teils gravierende Lücken auf. So geht nur eine Minderheit der Medien auf Probleme im Zusammenhang mit einem möglichen ›algorithmic bias‹ ein. Ob das Erfordernis, die Güte von Trainingsdaten sowie des Algorithmus gerade bei einer journalistischen Nutzung einer kritischen Überprüfung zu unterziehen, bei allen untersuchten Redaktionen vorhanden ist, lässt sich aus den Leitlinien nicht durchweg herauslesen. Die Untersuchung macht deutlich, dass es mit einer fortschreitenden Ausbreitung von KI in den Journalismus weiterer Untersuchungen bedarf. Sollten sich mehr Medienorganisationen dafür entscheiden, KI-Richtlinien auszuarbeiten und diese zu veröffentlichen oder der Forschung zugänglich zu machen, könnte eine größere Stichprobe weitere wichtige Punkte adressieren. Dazu zählt unter anderem die Frage, inwiefern sich die Regelungsbreite und -tiefe zwischen unterschiedlichen Organisationsformen unterscheidet und inwiefern mögliche professionelle und kulturelle Unterschiede zwischen den Redaktionen bei der Ausformulierung der Leitlinien in unterschiedlichen Ländern zutage treten.

Über den Autor

Kim Björn Becker, Dr. (*1986) ist seit 2018 Politischer Redakteur bei der *Frankfurter Allgemeinen Zeitung*. Zuvor arbeitete er bei der *Süddeutschen Zeitung* in München. Er unterrichtet journalistische Darstellungsformen an den Universitäten Trier und Mainz sowie an der Hochschule Darmstadt. Sein Forschungsschwerpunkt ist die Anwendung von Künstlicher Intelligenz im modernen Newsroom. Kontakt: kbb@kimbjoernbecker.com

Literatur

- ANP (2023): Leidraad: zo gaat de ANP-redactie om met AI. <https://twitter.com/CoolsHannes/status/1646115524235886592> (21.4.2023)
- AP (o.J.): *Leveraging AI to advance the power of facts. Artificial intelligence at The Associated Press.* <https://www.ap.org/discover/artificial-intelligence> (21.4.2023)

- BECKETT, CHARLIE (2019): *New powers, new responsibilities. A global survey of journalism and artificial intelligence*. <https://drive.google.com/file/d/1utmAMCmd4rfjHrUfLLfSJ-clpFTjyefi/view> (21.4.2023)
- BR (2020): *Unsere KI-Richtlinien im Bayerischen Rundfunk*. <https://www.br.de/extra/ai-automation-lab/ki-ethik-100.html> (21.4.2023)
- BBC (2023): *The BBC privacy promise*. <https://www.bbc.co.uk/usingthebbc/privacy/privacy-promise/> (22.4.2023)
- BBC (2021): *Responsible AI at the BBC: Our Machine Learning Engine Principles*. https://downloads.bbc.co.uk/rd/pubs/MLEP_Doc_2.1.pdf (21.4.2023)
- BUXMANN, PETER; SCHMIDT, HOLGER (2022): KI im Journalismus: Unterschätzter Helfer im Hintergrund. In: *Frankfurter Allgemeine Zeitung*. <https://www.faz.net/podcasts/f-a-z-kuenstliche-intelligenz-podcast/ki-im-journalismus-unterschaetzter-helfer-im-hintergrund-18218234.html> (21.4.2023)
- DEUTSCHER ETHIKRAT (2023): *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, Stellungnahme*. [https://www.ethikrat.org/mitteilungen/mitteilungen/2023/ethikrat-kuenstliche-intelligenz-darf-menschliche-entfaltung-nicht-vermindern/?cookieLevel=not-set#:~:text=pdf%20%7C%20104%20KB\)-,Stellungnahme,-\(pdf%20%7C%203%20MB](https://www.ethikrat.org/mitteilungen/mitteilungen/2023/ethikrat-kuenstliche-intelligenz-darf-menschliche-entfaltung-nicht-vermindern/?cookieLevel=not-set#:~:text=pdf%20%7C%20104%20KB)-,Stellungnahme,-(pdf%20%7C%203%20MB) (22.4.2023)
- DEUTSCHER JOURNALISTENVERBAND (2023): *Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus*. https://www.djv.de/fileadmin/user_upload/INFOS/Themen/Medienpolitik/DJV-Positionspapier_KI_2023-04.pdf (28.04.2023)
- DIAKOPOULOS, NICHOLAS (2019): *Automating the News. How Algorithms are rewriting the Media*. Cambridge/London: Harvard University Press.
- DIGNUM, VIRGINIA (2019): *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*. Cham: Springer Nature.
- DPA (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/> (21.4.2023)
- GOOGLETRENDS (2023): *Künstliche Intelligenz*. <https://trends.google.com/trends/explore?q=k%C3%BCnstliche%20intelligenz> (21.4.2023)
- MARCONI, FRANCESCO (2020): *Newsmakers. Artificial Intelligence and the Future of Journalism*. New York/Chichester: Columbia University Press.
- MENN, ANDREAS (2023): 100 Millionen Nutzer in zwei Monaten: Diese Grafiken zeigen den Hype um ChatGPT. In: *Wirtschaftswoche* vom 7. März 2023. <https://www.wiwo.de/my/technologie/digitale-welt/kuenstliche-intelligenz-100-millionen-nutzer-in-zwei-monaten-diese-grafiken-zeigen-den-hype-um-chatgpt/29019970.html?ticket=ST-295551-wFHWZOj97E2u9eukgqW-caso1.example.org> (21.4.2023)

- NATURE (2023): Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. In: *Nature* 613, 612 (2023), DOI: <https://doi.org/10.1038/d41586-023-00191-1>
- PNAS (2023): *The PNAS Journals Outline Their Policies for ChatGPT and Generative AI*, <https://www.pnas.org/post/update/pnas-policy-for-chatgpt-generative-ai> (21.4.2023)
- SHNEIDERMAN, BEN (2022): *Human-Centered AI*. Oxford: Oxford University Press.
- SIMON, FELIX (2022): Uneasy Bedfellows: AI in the News, Platform Companies and the Issue of Journalistic Autonomy. In: *Digital Journalism*, 10(10), S. 1832-1854. DOI: 10.1080/21670811.2022.2063150
- THOMSON REUTERS (2018): *Thomson Reuters Founders Share Company Limited*. <https://www.thomsonreuters.com/content/dam/ewp-m/documents/thomsonreuters/en/pdf/corporate-responsibility/thomson-reuters-founders-share-company-limited-1218.pdf> (21.4.2023)
- THOMSON REUTERS (o.J.): *Our AI principles*. <https://www.thomsonreuters.com/en/artificial-intelligence/ai-principles.html> (21.4.2023)
- OPENAI (2022): *Introducing ChatGPT*. <https://openai.com/blog/chatgpt> (21.4.2023)
- RAAD VOOR DE JOURNALISTIEK (o.J.): Nieuwe richtlijn over het gebruik van artificiële intelligentie in de journalistiek. <https://www.rvdj.be/nieuws/nieuwe-richtlijn-over-het-gebruik-van-artificiele-intelligentie-de-journalistiek> (22.4.2023)
- SKRUBBELTRANG MAHNKE, MARTINA; KARLIN, SIMON (2023): »Dieser Artikel könnte Sie auch interessieren« Entwicklung demokratisch verantwortungsvoller Algorithmen in Nachrichtenmedien. Eine dänische Fallstudie. In: *Communicatio Socialis*, 56(1), S.49-62. DOI: 10.5771/0010-3497-2023-1-49.
- VENTURA POCINO, PATRÍCIA (2021): *Algorithms in the newsrooms. Challenges and recommendations for artificial intelligence with the ethical values of Journalism*. Publikation des Catalan Press Council. https://fcic.periodistes.cat/wp-content/uploads/2022/03/venglishDIGITAL_ALGORITMES-A-LES-REDACCIONS_ENG-1.pdf (22.4.2023)
- WATERSON, JIM (2022): BBC licence fee to be abolished in 2027 and funding frozen. In: *The Guardian* vom 16. Januar 2022. https://www.theguardian.com/media/2022/jan/16/bbc-licence-fee-to-be-abolished-in-2027-and-funding-frozen?utm_source=dlvr.it&utm_medium=twitter (21.4.2023)
- WIRED (2023): *How WIRED Will Use Generative AI Tools*. <https://www.wired.com/about/generative-ai-policy/> (21.4.2023)
- WOLFENAGEL, EVA (2023): Wenn die KI lügt: Grenzen von ChatGPT für den Journalismus. In: *Deutschlandfunk* vom 4. Januar 2023, 15.55 Uhr. <https://www.deutschlandfunk.de/wenn-die-ki-luegt-grenzen-von-chatgpt-fuer-den-journalismus-dlf-ce94895e-100.html> (22.4.2023)